

(12)特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関  
国際事務局(43) 国際公開日  
2004 年 10 月 14 日 (14.10.2004)

PCT

(10) 国際公開番号  
WO 2004/088631 A1

- (51) 国際特許分類<sup>7</sup>: G10H 1/40, G10G 3/04  
 (21) 国際出願番号: PCT/JP2004/003010  
 (22) 国際出願日: 2004 年 3 月 9 日 (09.03.2004)  
 (25) 国際出願の言語: 日本語  
 (26) 国際公開の言語: 日本語  
 (30) 優先権データ:  
 特願2003-094100 2003 年 3 月 31 日 (31.03.2003) JP  
 (71) 出願人 (米国を除く全ての指定国について): ソニー株式会社 (SONY CORPORATION) [JP/JP]; 〒1410001 東京都品川区北品川 6 丁目 7 番 3 5 号 Tokyo (JP).  
 (72) 発明者; および  
 (75) 発明者/出願人 (米国についてのみ): 白石 吾朗 (SHI-RAISHI, Goro) [JP/JP]; 〒1410001 東京都品川区北品

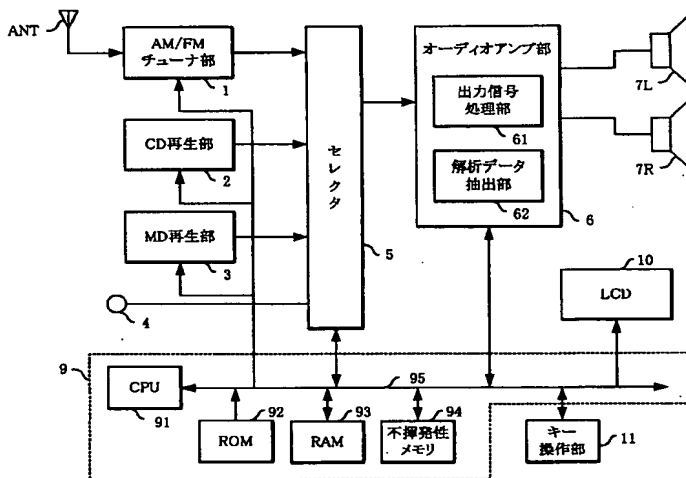
川 6 丁目 7 番 3 5 号 ソニー株式会社内 Tokyo (JP).  
 関根 千絵 (SEKINE, Chie) [JP/JP]; 〒1410001 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社内 Tokyo (JP). 増田 九美子 (MASUDA, Kumiko) [JP/JP]; 〒1410001 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社内 Tokyo (JP). 森 邦晴 (MORI, Kuniharu) [JP/JP]; 〒1410001 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社内 Tokyo (JP).

- (74) 代理人: 小池 晃, 外 (KOIKE, Akira et al.); 〒1000011 東京都千代田区内幸町一丁目 1 番 7 号 大和生命ビル 1 1 階 Tokyo (JP).  
 (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT,

[続葉有]

(54) Title: TEMPO ANALYSIS DEVICE AND TEMPO ANALYSIS METHOD

(54) 発明の名称: テンポ解析装置及びテンポ解析方法



- 1...AM/FM TUNER SECTION  
 2...CD REPRODUCTION SECTION  
 3...MD REPRODUCTION SECTION  
 5...SELECTOR  
 6...AUDIO AMPLIFICATION SECTION  
 61...OUTPUT SIGNAL PROCESSING SECTION  
 62...ANALYSIS DATA EXTRACTION SECTION  
 94...NON-VOLATILE MEMORY  
 11...KEY OPERATION SECTION

(57) Abstract: There is provided a tempo analysis device for analyzing the tempo of a musical composition. According to the level information contained in a sound signal from an analysis data extraction section (62), a control section (9) detects a peak position (level change apex) above a predetermined level by using a frame of a predetermined unit time section as a processing unit, calculates the interval between the peak positions (peak interval) in the frame duration, and decides the peak interval occurring with a high frequency as a tempo.

(57) 要約: 本発明は、楽曲などの音声のテンポを解析するテンポ解析装置であり、解析データ抽出部 (62) からの音声信号のレベル情報に基づいて、制御部 (9) において、所定の単位時間区間であるフレームを処理単位とし、所定レベル以上におけるピーク位置 (レベル変化の頂点) を検出して、当該フレーム区間における各ピーク位置間の間隔 (ピーク間隔) を求め、発生頻度の高いピーク間隔をテンポとして決定する。



LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI,  
NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG,  
SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ,  
VC, VN, YU, ZA, ZM, ZW.

NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG,  
CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

- (84) 指定国 (表示のない限り、全ての種類の広域保護が  
可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL,  
SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG,  
KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY,  
CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC,

添付公開書類:

— 国際調査報告書

2文字コード及び他の略語については、定期発行される  
各PCTガゼットの巻頭に掲載されている「コードと略語  
のガイダンスノート」を参照。

## 明細書

### テンポ解析装置及びテンポ解析方法

#### 技術分野

本発明は、楽曲などの音声信号から楽曲が演奏される速さであるテンポを抽出して利用できるようにするテンポ解析装置及び解析方法に関する。

本出願は、日本国において2003年3月31日に出願された日本特許出願番号2003-094100を基礎として優先権を主張するものであり、この出願は参照することにより、本出願に援用される。

#### 背景技術

従来、楽曲の音声データを解析することにより、その楽曲のテンポを自動的に抽出し、この抽出したテンポを、例えば、譜面を作成する際に利用したり、編曲をする際に利用したりするなどのことが行われている。この種の楽曲のテンポを抽出する技術の1つとして、特開2002-116754号公報に記載されたものがある。

この特許文献に記載された技術は、楽曲の音声データを時系列データとして取り込み、この音声データの自己相関を算出することにより当該音声データのピーク位置を検出して、テンポの候補を取得するようにし、一方で、自己相関パターンのピーク位置とそのレベルから当該楽曲のビート構造を解析し、テンポの候補とビート構造の解析結果とに基づいて、最も適切と思われるテンポを推定するものである。

この特許文献に記載された技術を利用することにより、音楽に対する先見的な知識を持っていなくても、誰でもが比較的簡単に、しかも正確に、目的とする楽曲のテンポを抽出し、これを利用することができるようになる。

ところで、最近、車載用オーディオシステム（カーステレオシステム）や家庭

用オーディオシステムにおいても、再生する楽曲のテンポを検出し、そのテンポに応じた情報を提供したり、あるいは、検出したテンポに応じて、種々の制御を行うようにすることが提案されている。

上述した特許文献に記載された技術は、音声データについて自己相関を算出したり、ビート構造を解析したりするなど、演算処理が複雑かつ膨大となり、実際に演算処理を行うCPU（Central Processing Unit）にかかる負担が大きくなる。

このため、上述した特許文献に記載の技術は、規模の比較的小さな車載用オーディオシステムや家庭用オーディオシステムに適用するには不向きな場合がある。また、上述した特許文献に記載の技術を用いようとする場合には、処理能力の高いCPUを用いたり、メモリ容量を大きくしたりする必要が生じるなど、オーディオシステムのコストアップにつながる可能性がある。

## 発明の開示

本発明の目的は、上述したような従来の技術が有する問題点を解決することができる新規なテンポ解析装置及びテンポ解析方法を提供することにある。

本発明の他の目的は、CPUに大きな負荷をかけることもなく、また、コストアップも生じさせないようにして、楽曲等の音声のテンポを簡単かつ正確に検出して利用できるテンポ解析装置及びテンポ解析方法を提供することにある。

上述したような目的を達成するために提案されるテンポ解析装置は、入力音声信号のレベル変化のピークのうち、所定の閾値より大きい複数のピークの位置を検出するピーク検出手段と、所定の単位時間区間において、ピーク検出手段により検出されるピーク位置の間の時間間隔を検出する間隔検出手段と、間隔検出手段により検出される時間間隔のうち、発生頻度の多い前記時間間隔に基づいて、前記音声信号により再生される音声のテンポを特定する特定手段とを備える。

本発明に係るテンポ解析装置は、ピーク検出手段により、音声信号のレベルについて、閾値より大きく、かつ、上昇から下降に転じる直前のピーク位置（レベル変化の頂点）が順次に検出される。そして、時間間隔検出手段により、所定の単位時間区間において検出される、一般的には複数個のピーク位置について、少

なくとも所定の1つのピーク位置を基準とし、このピーク位置とその他のピーク位置との時間間隔（ピーク間隔）が検出される。この後、特定手段により、時間間隔検出手段からの検出結果に基づいて、発生頻度の多い時間間隔が検出され、その時間間隔に基づいて、処理対象の音声信号により再生される楽曲などの音声のテンポが特定される。これにより、自己相関演算などの複雑な演算処理を行うことなく、簡単かつ正確に楽曲などの音声のテンポを特定することができる。

本発明に係るテンポ解析装置を構成する特定手段は、さらに具体的に、複数の単位時間区間において検出されるピーク位置間の時間間隔の発生頻度を累積し、この累積した発生頻度に基づいて、再生される音声の前記テンポを特定する。

本発明に係るテンポ解析装置は、さらに、入力信号を複数の周波数帯域に分離する帯域分離手段を備え、ピーク検出手段は、帯域分離手段により分離された複数の帯域のうちの少なくとも1つ以上の帯域毎に前記ピーク位置を検出するものであり、間隔検出手段は、ピーク検出手段により検出される少なくとも1つ以上の帯域毎のピーク位置の時間間隔を検出するものであり、特定手段は、少なくとも1つ以上の帯域毎に検出される時間間隔のうち、発生頻度の多い時間間隔に基づいて再生される音声の前記テンポを特定する。

また、本発明に係るテンポ解析装置は、音声信号の音量を算出する音量算出手段と、音量算出手段により算出された音量を基準として、ピーク位置を検出する場合に用いる前記閾値を設定するようにする閾値設定手段とを備える。

このテンポ解析装置において、帯域分離手段により分離された複数の帯域のうちの少なくとも1つ以上の帯域の音声信号の音量を算出する音量算出手段と、音量算出手段により算出された音量を基準として、ピーク位置を検出する場合に用いる閾値を設定するようにした閾値設定手段とを設けるようにしてもよい。

本発明に係るテンポ解析装置は、さらに、入力音声信号から所定の周波数帯域の音声信号を抽出する帯域抽出手段を備え、ピーク検出手段は、帯域抽出手段により抽出された音声信号についてピーク位置を検出するように構成してもよい。このテンポ解析装置において、帯域抽出手段で抽出された音声信号の音量を算出する音量算出手段と、音量算出手段により算出された音量を基準として、ピーク位置を検出する場合に用いる閾値を設定する閾値設定手段を設けるようにする。

本発明に係るテンポ解析装置は、さらに、画像表示素子と、画像表示素子に表示可能な複数の画像の画像データを記憶する記憶手段と、特定手段により特定される前記テンポに基づいて、前記記憶手段から画像データを選択して読み出し、読み出した前記画像データに応じた画像を前記画像表示素子に表示する表示制御手段とを備える。

このテンポ解析装置の表示手段は、記憶手段から読み出す画像データに応じた画像を画像表示素子に表示する画像の大きさ、移動速度、移動パターンの少なくとも1つを制御する。

また、表示手段は、特定手段により特定されるテンポと音量算出手段により算出された音量とに基づいて、記憶手段から画像データを選択して読み出すようにしてもよい。

そして、本発明に係るテンポ解析方法は、入力音声信号のレベル変化のうち、所定の閾値より大きい複数のピークの位置を検出し、所定の単位時間区間において、検出した前記ピーク位置の間の時間間隔を検出し、検出した前記時間間隔のうち、発生頻度の多い時間間隔に基づいて、入力音声信号により再生される音声のテンポを特定する。テンポの特定に際しては、複数の前記単位時間区間において検出されるピーク位置間の時間間隔の発生頻度を累積し、この累積した発生頻度に基づいて、再生される音声の前記テンポを特定する。

本発明に係るテンポ解析方法は、さらに、入力音声信号を複数の周波数帯域に分離し、ピーク位置の検出に際しては、分離された前記複数の周波数帯域の少なくとも1つ以上の帯域毎に前記ピーク位置を検出し、時間間隔の検出に際しては、少なくとも1つ以上の前記帯域毎のピーク位置の時間間隔を検出し、テンポの特定に際しては、少なくとも1つ以上の帯域毎に検出される時間間隔のうち、発生頻度の多い時間間隔に基づいて、再生される音声のテンポを特定する。

また、本発明に係るテンポ解析方法は、入力音声信号から所定の周波数帯域の音声信号を抽出し、ピーク位置を検出に際しては、抽出された音声信号についてのピーク位置を検出するようにしてもよい。

さらに、本発明に係るテンポ解析方法は、入力音声信号の音量を算出し、算出した音量を基準として、ピーク位置を検出する場合に用いる閾値を設定するよう

にしてもよい。

本発明に係るテンポ解析方法は、特定されたテンポに基づいて、記憶手段に記憶されている複数の画像データの中から画像データを選択して読み出し、読み出した前記画像データに応じた画像を画像表示素子に表示する。このテンポ解析方法は、特定されたテンポに基づいて、画像表示素子に表示する画像の大きさ、移動速度、移動パターンを制御する。又は、特定されたテンポと算出された音量とに基づいて、記憶手段に記憶されている複数の画像データを選択して読み出すようにしてもよい。

本発明の更に他の目的、本発明によって得られる具体的な利点は、以下において図面を参照して説明される実施の形態の説明から一層明らかにされるであろう。

#### 図面の簡単な説明

図 1 は、本発明を適用したカーステレオ装置を示すブロック図である。

図 2 は、カーステレオ装置に搭載されるテンポ分析装置を示すブロック図である。

図 3 は、制御部で実行されるメイン処理を説明するためのフローチャートである。

図 4 は、図 3 に示すステップ S 1 において実行される総音量計算処理を説明するためのフローチャートである。

図 5 は、図 3 に示すステップ S 2 において実行されるテンポ抽出処理を説明するためのフローチャートである。

図 6 は、図 5 に示すステップ S 2 1 において実行されるスレッシュホールド処理を説明するためのフローチャートである。

図 7 は、図 5 に示すステップ S 2 3 において実行されるピーク位置抽出処理を説明するためのフローチャートである。

図 8 は、ピーク位置抽出処理を説明するための図である。

図 9 は、図 5 に示すステップ S 2 5 において実行されるピーク間隔（周期）リスト作成及びテンポ決定処理を説明するためのフローチャートである。

図 1 0 は、周期リスト（ピーク間隔リスト）を説明するための図である。

図 1 1 は、周期リストの足切り処理を説明するための図である。

図 1 2 は、各フレーム毎の発生頻度が最も高いピーク間隔の保持と利用について説明するための図である。

図 1 3 は、決定されたテンポと音量とにより利用可能な画像データが特定される構造について説明するための図である。

図 1 4 は、決定されたテンポを用いて選択され表示するようにされる画像の表示例を示す図である。

発明を実施するための最良の形態

以下、本発明に係るテンポ解析装置及びテンポ解析方法を図面を参照しながら説明する。

なお、以下に説明では、本発明をカーステレオ装置（カーオーディオシステム）に適用した例をあげて説明する。

まず、本発明に係るカーステレオ装置を説明する。本発明が適用されるカーステレオ装置は、図 1 に示すように、ラジオ放送の受信アンテナ ANT、AM/FM チューナ部 1、CD（Compact Disc）再生部 2、MD（Mini Disc）再生部 3、外部接続端子 4、入力セレクト 5、オーディオアンプ部 6、左右のスピーカ 7 L、7 R、制御部 9、LCD（Liquid Crystal Display）10、キー操作部 11 を備える。

制御部 9 は、図 1 に示すように、CPU（Central Processing Unit）91、ROM（Read Only Memory）92、RAM（Random Access Memory）93、不揮発性メモリ 94 が CPU バス 95 により接続されて形成されたマイクロコンピュータであり、このカーステレオ装置の各部を制御する。

ここで、ROM 92 は、CPU 91 によって実行されるプログラムや処理に必要なデータ、表示に用いる画像データや文字フォントデータなどが記憶されたものである。RAM 93 は、主に作業領域として用いられる。不揮発性メモリ 94 は、例えば、EEPROM（Electrically Erasable and Programmable RO



M) やフラッシュメモリであり、当該カーステレオ装置の電源が落とされても保持しておく必要のあるデータ、例えば、各種の設定パラメータなどを記憶保持する。

また、制御部 9 には、図 1 に示すように、LCD 10 と、キー操作部 11 とが接続されている。LCD 10 は、比較的に大きな表示画面を有するものであり、このカーステレオ装置の状態や操作ガイダンスなどを表示することができるとともに、例えば、外部入力端子を通じて、GPS (Global Positioning System) や DVD (Digital Versatile Disc) の再生装置が接続された場合には、制御部 9 の制御によって、地図情報や動画情報等を表示する。

キー操作部 11 は、各種の操作キーやファンクションキー、操作ダイヤルなどを備えたものであり、ユーザからの操作入力を受け付けて、これを電気信号に変換し、制御部 9 に通知することができる。これにより、制御部 9 は、ユーザからの指示に応じて、このカーステレオ装置の各部を制御するようにしている。

そして、図 1 に示したように、このカーステレオ装置は、音声信号（音声データ）等の供給端として、AM/FM チューナ部 1、CD 再生部 2、MD 再生部 3、外部入力端子 4 を備える。AM/FM チューナ部 1 は、制御部 9 からの選局制御信号に基づいて、AM ラジオ放送又は FM ラジオ放送のうちの目的とする放送チャンネルを受信、選局し、この受信、選局したラジオ放送信号を復調して、復調後の音声信号をセクタ 5 に供給する。

CD 再生部 2 は、スピンドルモータ、光学ヘッド部などを備え、これに装填された CD を回転駆動し、当該 CD にレーザ光を照射して、その反射光を受光することにより、CD に微少な凹凸の連続であるピットパターンとして記録されている音声データを読み出す。そして、読み出した音声データを電気信号に変換し、復調して再生用の音声信号を形成し、これをセクタ 5 に供給する。

MD 再生部 3 は、CD 再生部 2 の場合と同様に、スピンドルモータ、光学ヘッド部などを備え、これに装填された MD を回転駆動し、当該 MD にレーザ光を照射して、その反射光を受光することにより、当該 MD に磁化変化として記録されている音声データを読み出し、これを電気信号に変換する。読み出された音声データは、通常、データ圧縮されているので、これをデータ伸張処理（圧縮解凍処理）して再生用の音声信号を形成し、これをセクタ 5 に供給する。

また、外部接続端子 4 には、上述したように、例えば G P S や D V D 再生装置などの外部機器が接続され、それらの機器からの音声信号が、セレクト 5 に供給するようにされる。

そして、セレクト 5 は、制御部 9 により切り換え制御がなされ、A M / F M チューナ 1、C D 再生部 2、M D 再生部 3、外部入力端子 4 のうちのいずれの部分からの音声信号を出力するかを切り換える。これにより、A M / F M チューナ 1、C D 再生部 2、M D 再生部 3、外部入力端子 4 のうちの目的とする部分からの音声信号がオーディオアンプ部 6 に供給される。

オーディオアンプ部 6 は、大きく分けると、出力信号処理部 6 1 と解析データ処理部 6 2 とからなっている。出力信号処理部 6 1 は、制御部 9 からの制御信号に基づいて、出力しようとする音声信号についての音量調整、音質調整等の各種の調整処理を行って、出力用の音声信号を形成し、これをスピーカ 7 L、7 R に供給する。

これにより、図 1 において参照符号 1 から 4 で示した部分の内の目的とする供給部分からの音声信号に応じた音声をスピーカ 7 L、7 R から放音することができるようにされる。

一方、解析データ抽出部 6 2 は、これに供給された音声信号を複数の周波数帯域に分割し、各周波数帯域の音声信号のレベルを示す情報を制御部 9 に供給する。制御部 9 は、詳しくは後述するが、解析データ抽出部 6 2 からの解析データに基づいて、音声信号のピーク位置を検出し、所定単位時間におけるピーク位置間の時間間隔を算出して、この算出結果に基づいて、出力する音声のテンポを特定する。

そして、本例の制御部 9 は、例えば、R O M 9 2、あるいは、不揮発性メモリ 9 4 に記憶されている静止画像データの中から、上述のようにして特定したテンポに応じたものを選択し、それを L C D 1 0 に表示するようにしている。また、制御部 9 は、L C D 1 0 に表示するようにした静止画像に重ねて、例えば、図形やキャラクタなどの画像を、特定したテンポに応じて動くような態様で表示を行うようにしている。

このように、本発明に係るカーステレオ装置においては、オーディオアンプ部

6の解析データ抽出部62と制御部9とによりテンポ解析装置を構成し、これらが協働することによって、再生する楽曲などの音声のテンポを特定して、これを利用することができるようにしている。

すなわち、解析データ抽出部62と制御部9とにより構成されるテンポ解析装置部が、本発明に係るテンポ解析装置の一実施の形態が適用されたものであり、ここで用いられる方法が、この発明に係るテンポ解析方法の一実施の形態が適用されたものである。

そして、本発明においては、以下に詳述するように、再生しようとする楽曲などの音声のテンポを特定する際には、従来のように自己相関算出等の複雑な演算処理を行うことはなく、簡単な処理で、しかも正確に目的とする音声のテンポを特定するようにしている。

次に、本発明に係るカーステレオ装置に搭載されたテンポ解析装置部について説明する。

図2は、このカーステレオ装置に搭載されたテンポ解析装置部を示すブロック図である。上述したように、本発明に係るテンポ解析装置は、カーステレオ装置のオーディオアンプ部6に設けられる解析データ抽出部62と、制御部9とにより構成される。

図2に示すように、解析データ抽出部62と制御部9との間には、A/D変換部12が設けられる。このA/D変換部12は、解析データ抽出部62から出力される音声信号のレベルを示す情報（例えば電圧値）を例えば、0～1023までの1024ステップのデジタルデータに変換して制御部9に供給するようにするものである。

このA/D変換部12は、図2に示したように、解析データ抽出部62と制御部9との間に設けることも可能であるが、解析データ抽出部62の機能として設けるようにすることもできるし、また、制御部9の機能として設けるようにすることもできる。

この実施の形態において、解析データ抽出部62は、これに供給された音声信号を複数の周波数帯域に分離する帯域分離部621と、複数の周波数帯域に分離された音声信号のそれぞれのレベルを検出し、これをレベル情報として出力する

レベル検出部 6 2 2 とからなっている。

帯域分離部 6 2 1 は、図 2 にも示したように、中心周波数が、6 2 H z、1 5 7 H z、3 9 6 H z、1 k H z、2 . 5 1 k H z、6 . 3 4 k H z、1 6 k H z の 7 つの周波数帯域（7 バンド）に分離するようにしている。

帯域分離部 6 2 1 において、各周波数帯域に分離された音声信号のそれぞれは、図 2 に示したように、レベル検出部 6 2 2 に供給され、そのそれぞれごとにレベルが検出される。レベル検出部 6 2 2 において検出された各周波数帯域の音声信号のレベルを示す情報は、A/D 変換部 1 2 を通じて制御部 9 に供給される。すなわち、帯域分割された各帯域の音声信号のレベル波形（音声レベル波形）がデジタルデータとして制御部 9 に供給される。

なお、解析データ抽出部 6 2 は、汎用の集積回路、例えば、I C A 6 3 3 A B (STMicroelectronics) 等を用いて実現することが可能である。また、解析データ抽出部 6 2 をマイクロコンピュータで構成するようにし、ここで実行されるソフトウェアによって音声信号の帯域分割や信号レベルの検出を行うようにすることもできる。

そして、制御部 9 は、解析データ抽出部 6 2 からの各周波数帯域の音声信号のレベル（音声レベル波形）を用い、ごく簡単な比較処理を中心とする処理により、処理対象の音声のテンポを特定する。そして、特定したテンポに基づいて、制御部 9 は、例えば R O M 9 2 に用意された静止画像データの中からそのテンポに応じた静止画像を形成する画像データを抽出し、それを L C D 1 0 の表示画面に表示するようにする。

同時に、制御部 9 は、所定の図形やキャラクタなどを L C D 1 0 の表示画面に表示するようにするとともに、その図形やキャラクタを、特定したテンポに応じて移動させるようにしたりするなどのことを行うようにしている。

次に、上述したように、制御部 9 の機能として行われる処理対象の音声信号により再生される音声のテンポを特定する処理について具体的に説明する。図 3 は、本発明に係るカーステレオ装置において行われる処理対象の音声信号により再生される音声のテンポを特定する場合の処理手順を示すフローチャートである。

このカーステレオ装置において、制御部 9 は、まず、最終的に特定されたテン

ボとともに画像データの表示のためのパラメータとなる入力音声信号の音量レベル（総音量）の計算処理を行う（ステップS1）。

次に、制御部9は、処理対象の音声についてのテンポの抽出及び特定のための処理を行う（ステップS2）。このステップS1、ステップS2の処理によって求められたパラメータ（総音量とテンポ）により、表示する画像データや表示内容が決定される。

そして、本発明に係るカーステレオ装置においては、上述したように処理対象の音声信号を7つの周波数帯域（7バンド）に分割し、所定の時間単位区間（1フレーム）を処理単位として処理を行うようにしている。ここで、時間単位区間（1フレーム）は、連続する例えば4秒間の区間である。

そして、1フレーム（4秒間）の区間をサンプリング周波数が20Hzのクロック信号を用いてサンプリングすることにより、1フレームに80サンプルを得るようにしている。さらに、例えば、10フレーム、20フレームなどのように、所定のフレーム数分の情報を累積し、この累積した情報に基づいて、総音量の算出やテンポの決定（特定）を行うようにしている。

次に、図3に示す処理のステップS1の処理、及びステップS2の処理の詳細について説明する。

まず、ステップS1の総音量の計算処理について説明する。図4は、図3に示すステップS1において行われる処理を説明するためのフローチャートである。

ここでは、図4にも示すように、処理結果を累積する連続した複数フレームの各フレームにおける7バンドの合計音量のデータバッファをVolData[Frame]とし、各バンド毎の音量データ（レベルデータ）の格納バッファをdata[band]とし、総音量の値の格納バッファをTotalVolとする。

また、[Frame]は、総音量の計算対象となるフレーム数であり、[Frame]番目に相当するフレームは、処理結果を累積する連続した複数フレームの内の最古のフレームである。[band]は、いずれのバンド（周波数帯域）かを示すバンド番号でもある。

そして、現在処理の対象となっている最新のフレームの音量バッファをVolData[1]とし、処理結果を累積する連続した複数フレームの内の最古のフレ

ームの音量バッファをV o l D a t a [F r a m e]とすると、図4に示すように、制御部9のCPU91は、まず、総音量T o t a l V o lから、最古のフレームの音量を減算する（ステップS11）。

次に、バッファV o l D a t a [1]～V o l D a t a [F r a m e]に格納データを、1バッファずつシフトする（ステップS12）。例えば、V o l D a t a [F r a m e] = V o l D a t a [5]である場合を例にすると、V o l D a t a [4]のデータをV o l D a t a [5]にシフトし、V o l D a t a [3]のデータをV o l D a t a [4]にシフトし、V o l D a t a [2]のデータをV o l D a t a [3]にシフトし、V o l D a t a [1]のデータをV o l D a t a [2]にシフトすることになる。

そして、解析データ抽出部62からの最新のフレームの各バンド（周波数帯域）のレベルデータd a t a [1]、d a t a [2]、d a t a [3]、d a t a [4]、d a t a [5]、d a t a [6]、d a t a [7]を合算し、この合算結果を最新のフレームの音量を示すデータとして、バッファV o l D a t a [1]にセットする（ステップS13）。

そして、ステップS13において求めた、最新の処理対象のフレームの音量の値を総音量の値を保持するT o t a l V o lの値に加算することにより、最新のフレームから過去にさかのぼる方向に総音量を計算する[F r a m e]分のフレームを対象とした総音量が求められる（ステップS14）。

このようにして、処理対象の音声信号の総音量が算出され、この算出された総音量をパラメータのひとつとして用いることにより、画像データを選択・表示することができるようにされる。

なお、上述した総音量の計算処理は、複数の周波数帯域に分割された音声レベル波形から求めるようにしたが、これ以外に、供給された音声信号に対する音声レベル波形から求めてもよいし、例えば中音域のような特定の周波数帯域成分を取り出すフィルタを用意してその帯域の音声信号に対する音声レベル波形から求めるようにしてもよい。

次に、図3に示したステップS2において行われるテンポ抽出処理について具体的に説明する。図5は、図3に示したステップS2において行われるテンポ抽

出処理を説明するためのフローチャートである。図5に示すように、ステップS21からステップS24までの各処理は、帯域分割された各バンド毎の音声信号を対象として行われる。

すなわち、制御部9のCPU91は、各バンド毎に、スレッシュホールドを設定する処理を行い（ステップS21）、例えばRAM93、あるいは、不揮発性メモリ94に設けられるピーク位置検出用のバッファであるピークバッファの内容のシフト処理を実行する（ステップS22）。そして、ステップS21で設定したスレッシュホールド以上のレベルのピーク位置（レベル変化の頂点）を抽出する処理を行い（ステップS23）、抽出したピーク位置に基づいて、各ピーク位置間のピーク間隔（ピーク位置間の時間間隔）を求める（ステップS24）。

各バンド（帯域）毎に行われるステップS21～ステップS24までの処理の後、制御部9のCPU91は、各バンド毎のピーク間隔を1つのリストにまとめる処理を行い、検出頻度（発生頻度）の最も高いピーク間隔（ピーク周期）を再生している音声のテンポとして特定する（ステップS25）。

次に、図5に示したテンポ抽出処理のステップS21のスレッシュホールド処理、ステップS23のピーク抽出処理、ステップS25のテンポを特定する処理のそれぞれについてより詳細に説明する。

図6は、図5に示したテンポ抽出処理のステップS21において行われるスレッシュホールド処理を説明するためのフローチャートである。この実施の形態においては、図3に示したステップS1において実行される処理に類似する処理であって、帯域分割された各バンド毎に1フレーム（4秒間）の区間にわたりそれぞれの最大音量レベルを求め、その値をMaxVol[band]として保持しておく。次の1フレーム（4秒間）の区間に対してスレッシュホールド処理を行う際に、保持されてあるMaxVol[band]を呼び出して、この値に、例えば0.8を掛け算することにより、最大音量MaxVol[band]の80パーセントのレベルを求め、この求めたレベルが前の1フレーム（4秒間）の区間に対して求められたスレッシュホールドThresより大きいかな否かを判断する（ステップS211）。

ステップS211の判断処理において、スレッシュホールドThresが、最大音

量 $MaxVol[band]$ の80パーセントのレベルよりも大きいと判断した場合には、音量が低下していると判断し、スレッシュホールド $Thres$ に、当該スレッシュホールド $Thres$ の90パーセントのレベルを設定するようにする（ステップ $S212$ ）。

ステップ $S211$ の判断処理において、スレッシュホールド $Thres$ が、音量 $MaxVol[band]$ の80パーセントのレベルよりも小さいと判断したときには、音量が上がっていると判断し、今回の新たな最大音量 $MaxVol[band]$ の80パーセントのレベルをスレッシュホールド $Thres$ に設定するようにする（ステップ $S213$ ）。

このように、本発明に係るカーステレオ装置においては、各バンド毎に音量が低下した場合と上昇した場合との両方において、スレッシュホールド $Thres$ を適切に変更することができるようにしている。このスレッシュホールド $Thres$ を、音声信号のピーク位置を検出する場合の基準として用いることによって、音声のテンポを正確に特定することができるようにしている。

次に、図5に示したテンポ抽出処理のステップ $S23$ において行われるピーク位置の抽出処理について説明する。図7は、図5に示したステップ $S23$ において実行されるピーク位置の抽出処理を説明するためのフローチャートである。上述もしたように、この実施の形態においては、サンプリング周波数が $20Hz$ のクロック信号を用い、音声信号は、1フレームである4秒間に80回サンプリングされて、そのレベルが検出するようにされる。そして、各サンプルについて、図7に示す処理が行われることになる。

まず、制御部9は、現在のサンプルのレベルが、図6を用いて説明したようにして設定されるスレッシュホールド $Thres$ を下回っているか否かを判断する（ステップ $S231$ ）。このステップ $S231$ の判断処理において、現在のサンプルのレベルが、スレッシュホールド $Thres$ を下回っていないと判断したときには、現在のサンプルのレベルが最大値である可能性があるので、既に最大値の候補として仮登録されているレベルと現在のサンプルのレベルとを比較し、現在のサンプルのレベルの方が高いか否かを判断する（ステップ $S232$ ）。

ステップ $S232$ の判断処理において、現在のサンプルのレベルよりも、既に



登録されている最大値の候補のレベルの方が高ければ、何もすることなく、この図7に示す処理を抜ける。ステップS232の判断処理において、現在のサンプルのレベルの方が、仮登録されている最大値の候補のレベルよりも高い場合には、現在のサンプルのレベルと当該サンプルの位置を仮登録し（ステップS233）、この図7に示す処理を抜ける。なお、仮登録は、例えば、RAM93、あるいは、不揮発性メモリ94の仮登録エリアにする。

また、ステップS231の判断処理において、現在のサンプルのレベルが、スレッシュホールドThresを下回っていると判断したときには、ステップS233において仮登録したレベルのサンプル位置は、現在の処理対象のフレーム内か否かを判断する（ステップS234）。

ステップS234の判断処理において、仮登録したレベルのサンプル位置は、現在の処理対象のフレーム内ではないと判断したときには、処理の対象となっているフレームが次のフレームに移っているので、何もすることなく、この図7に示す処理を抜けるようにする。

ステップS234の判断処理において、仮登録したレベルのサンプル位置は、現在の処理対象のフレーム内であると判断したときには、ピークの候補として仮登録したレベルとそのサンプリング位置とを、ピークレベル及びピーク位置として、所定のエリア（最大値位置情報エリア）に追加記録するとともに、ピークの数に1カウントし、この図7に示す処理を抜ける。

このように、本発明に係るカーステレオ装置においては、自己相関の算出を行うことなく、比較的簡単な比較処理だけで、ピークレベルを検出し、そのピークレベルの位置（ピーク位置）を抽出することができるようにしている。

そして、このカーステレオ装置においては、図7に示した処理が、図5に示した処理のステップS23において行われることにより得られるピーク位置に基づいて、図5に示したステップS24においては、ピーク間隔（ピーク位置間の時間間隔）が求められる。

図8は、本発明において行われるピーク間隔の検出処理を説明するための図である。図8に示すように、1フレーム内において、スレッシュホールドThres以上のピーク位置（ピーク点）が4つ存在する場合を例にして、ピーク間隔を求め

る処理について説明する。

制御部 9 は、例えば、R A M 9 3 あるいは不揮発性メモリに記憶保持されたピーク位置を示す情報に基づいて、図 8 において、アルファベット A、B、C、D、E、F が示すように、同じ区間が重複することがないように、ピーク間隔を求める。

図 8 に示した例では、4 つのピーク位置のそれぞれを基準にして、他のピーク位置との間隔を求めるようにする。しかし、基準となるピーク位置と他のピーク位置とが逆になるだけの区間は、区間の重複となるので、実質的に区間が重複する場合には、その一方だけを生かすように処理する。

したがって、図 8 に示した例の場合には、4 つのピーク位置のそれぞれについて、他の 3 つのピーク位置との間でピーク間隔が求められるので、12 個のピーク間隔を検出することができるが、重複する区間については、そのうちの 1 つしか生かさないようにすることによって、図 8 に示すように、6 つのピーク間隔 A、B、C、D、E、F が検出できる。

この処理は、処理対象のフレーム区間の各バンドのレベルデータを対象として行われる。そして、当該処理対象のフレーム区間の各バンドにおける求められたピーク間隔をピーク間隔（周期）リスト（以下、周期リストという。）に展開し、この周期リストに基づいて、再生するようにしている楽曲のテンポが決定（特定）するようにされる。

図 9 は、図 5 に示したステップ S 2 5 において実行される周期リスト作成及びテンポ決定処理を説明するためのフローチャートである。図 9 に示すフローチャートの処理は、制御部 9 において実行される処理である。

まず、制御部 9 は、現在、音量がゼロであるか否かを判断する（ステップ S 2 5 1）。この判断は、前述した総音量 T o t a l V o l をチェックすることにより行うこともできるし、また、別途に、入力音声信号についての音量レベルを検出し、これをチェックするようにしてもよい。

なお、音量が完全にゼロにならない場合もあることを想定し、ステップ S 2 5 1 の処理においては、例えば、規定スレッシュホールド以下の音声レベルの音声信号が規定サンプル以上続いた場合には、音量がゼロになった、すなわち、楽曲の再

生が終了したと判断するようにしてもよい。

ステップS 2 5 1の判断処理において、音量がゼロでないと判断したときには、制御部9は、図7を用いて前述したようにして求められるすべてのピーク間隔をスコアに重み付けをしながら周期リストに展開する（ステップS 2 5 2）。周期リストは、例えば、図10に示すように、横軸をピーク間隔、縦軸をスコア（検出数）として、処理対象のフレーム区間における各バンドにおいて検出した各ピーク間隔について、その検出回数を累積するようにするものである。

ここで、重み付けは、各バンド毎、ピーク間隔の大小により所定の値を予め設定しておく。例えば、高音域のバンドに対する重み付けを、中音域のバンドに対する重み付けよりも小さい値とするようにしてもよい。あるいは、各バンドに対する重み付けを同一の値とするようにしてもよい。

なお、この例においては、図10に示したように、各バンド毎の重み付けをW1、W2、W3、…で示し、ピーク間隔毎の重み付けをAA、BBで示している。ここでスコアの計算例は次のようになる。

ピーク間隔B、Eのスコア＝AA＊（1バンド目スコア＊W1＋2バンド目スコア＊W2＋・・・＋6バンド目スコア＊W6＋7バンド目スコア＊W7）

この例においては、ピーク間隔毎の重み付けと各バンド毎の重み付けとを行うことにより、各ピーク間隔のスコアを得るようにしている。

そして、図9に示した周期リストにおいては、図8を用いて説明したように検出されるピーク間隔の内、同じ間隔であるピーク間隔B、Eの検出回数が最も多く検出されていることがわかる。制御部9は、作成した周期リストから、検出回数、すなわち積み上げられたスコアの最も高いピーク間隔をテンポとして決定（特定）する（ステップS 2 5 3）。

次に、制御部9は、周期リストのスコアの最大値が予め決められた規定値を超えているか否かを判断する（ステップS 2 5 4）。テンポの決定は、周期リストに基づいて迅速に行わなければならないので、周期リストに必要以上のデータを蓄積することは、処理の遅延、メモリの無駄使い等につながる可能性があるため望ましくない。

ステップS 2 5 4の判断処理において、周期リストのスコアの最大値が予め決

められた規定値を超えていない場合には、図 9 に示す処理を終了する。また、ステップ S 2 5 4 の判断処理において、周期リストのスコアの最大値が予め決められた規定値を超えていると判断した場合には、周期リストのデータについての足切り処理を行い（ステップ S 2 5 5）、この後、この図 9 に示す処理を終了する。

ステップ S 2 5 5 において行われる周期リストの足切りは、上述もし、また、図 1 1 にも示すように、累積されていく各ピーク間隔のスコアが、規定値を超えた場合に行われる。具体的には、周期リストの各ピーク間隔のスコアから所定スコア分を減算するようにしたり、あるいは、周期リストに展開したデータのうち、例えば、一番古いフレームの各ピーク間隔のスコアを差し引くようにしたり、あるいは、一番古いフレームから新しいフレーム方向に複数フレーム分のピーク間隔のスコアを差し引くようにすることにより行われる。

また、図 9 に示したステップ S 2 5 1 の判断処理において、音量がゼロであると判断したときには、楽曲の再生が終わったと判断することができるので、図 1 0 に示したように作成される周期リストをリセットし（ステップ S 2 5 6）、新たに再生される楽曲のテンポの解析処理に備えるようにして、この図 9 に示す処理を終了する。

なお、このカーステレオ装置において、制御部 9 は、各フレームにおいて検出されるそのフレームにおける検出頻度の最も高いピーク間隔を示す情報が、複数フレーム分、例えば 1 0 0 0 フレーム分蓄積するようにされる。例えば、図 1 2 に示すように、各フレームの検出頻度の最も高いピーク間隔を示すデータが保持するようにされる。

このように、処理対象となった過去のフレームについても、ピーク間隔を示す情報を保持しておくことにより、例えば、あるフレームで突然ピーク間隔が大きく変わったような場合であっても、その前後のフレームのピーク間隔を示す情報を参照することによって、ピーク間隔の突然の変動に大きな影響を受けることなく、適正に再生対象の楽曲のテンポを決定することができるようになる。

そして、本発明に係るカーステレオ装置において、制御部 9 は、上述のようにして、再生対象の楽曲のテンポを決定すると、その決定したテンポに応じて、ROM 9 2 に保持されている例えば静止画像の画像データを読み出し、この読み出

した画像データによる静止画像をLCD10に表示するようにしている。

このカーステレオ装置において、LCD10に表示される静止画像は、再生している楽曲のテンポと音量とに基づいて決められる。すなわち、図13に示すように、横軸をテンポとし、縦軸を音量とする座標平面を想定し、この平面上に9ブロック×9ブロックの領域を設けるようにする。

そして、テンポと音量とにより決まるブロックに対応して、画像を形成する画像データが一意に決まるようにしている。つまり、図13に示した81個のブロックのそれぞれに対して、所定の画像を形成する画像データが決まるようにされている。

したがって、例えば、図13に示したように、テンポTPと、音量Vがわかれば、これで示される座標(TP, V)が属するブロックに割り当てられた画像データがROM92から読み出され、この読み出された画像データによる静止画像が、制御部9の制御によって、LCD10の表示画面に表示するようにされる。

なお、ここでは、例えば、ROM92には、少なくとも図13に示したように設定される81ブロックのそれぞれに対応する81枚の静止画像を形成する画像データが記憶保持される。しかし、実際には、図13に示したいずれのブロックにも属さない場合も生じる可能性があるので、どのブロックにも属さない場合に用いる静止画像を形成する複数の画像データをも記憶保持し、これを利用することもできるようにされる。したがって、例えばROM92は、この実施の形態の場合、100枚前後の静止画像の画像データが記憶保持されている。

また、本発明に係るカーステレオ装置においては、LCD10の表示画面に、テンポと音量とに応じた静止画像を表示するものとして説明したが、所定時間分の動画像を表示したり、所定時間分の動画を繰り返し表示するなど、動画像の表示を行うようにすることももちろん可能である。

さらに、本発明に係るカーステレオ装置においては、楽曲の再生時において、上述したようにテンポと音量とに応じた画像をLCD10の表示画面に表示するだけでなく、例えば、図14において、オブジェクトObが示すように、予め決められた図形やキャラクタなどの表示オブジェクトをLCD10の表示画面に表示させ、これを移動させるようにしている。

この場合、オブジェクトObの移動パターンや移動速度などは、例えば、決定されたテンポに応じて決められ、テンポが速ければ、激しく動かし、テンポが遅ければ、ゆっくりと動かすなどというように制御することになる。もちろん、テンポと音量とにより、移動パターンや移動速度を選択するようにしてもよい。また、表示して移動させるようにする表示オブジェクト自体についても複数個用意しておき、決定したテンポ、あるいは、決定したテンポと音量とによって、用いる表示オブジェクトを選択するようにすることもできる。

このように、本発明に係るカーステレオ装置においては、自己相関演算などの複雑な演算処理を行うことなく、再生する楽曲などの音声のテンポを簡単に、しかも迅速かつ正確に特定することができるようになる。したがって、カーステレオ装置の制御部に大きな負荷をかけることなく、再生する音声のテンポを特定することができる。

そして、特定したテンポに応じてLCD10に表示する画像を特定し、これを表示してユーザに提供することができるようになる。また、特定したテンポに応じて、表示オブジェクトをLCDの表示画面に表示させ、これをテンポに応じて移動させるなどのことができるようになる。つまり、物理的な情報を利用するグラフィックイコライザとは異なり、音楽的な情報である特定したテンポに応じて、画像情報を提供することができるという、新たな態様での情報の提供ができるようになる。

なお、上述した実施の形態においては、再生する音声信号を7つの周波数帯域に分割して、各帯域毎に処理するものとして説明したが、これに限るものではない。分割する周波数帯域数は、いくつでもよい。すなわち、必ずしも周波数帯域を分割する必要はなく、全周波数帯域を有する音声信号に対して上述した処理を行うようにしてももちろんよい。

また、処理対象の音声信号を複数の周波数帯域に分割するようにした場合であっても、その分割されたすべての周波数帯域の音声信号を処理対象とする必要はなく、分割した周波数帯域の1つ以上の帯域を選択して処理対象とするようにしてもよい。あるいは、バンドパスフィルタにより処理対象とする周波数帯域の音声信号を抽出して上述した処理を行うようにしてもよい。

また、ピーク位置の検出に際しては、音声波形のレベルについてのスレッシュホールドを、前フレーム区間の最大音量に基づいて算出するようにしたが、これに限るものではない。音声波形についてのスレッシュホールドは、所定の値を用いるように予め設定しておくことも可能である。また、選択された音量レベルなどに応じて、予め決められた複数の値の中から所定の値を選択してこれを用いるようにしてもよい。

前述した実施の形態においては、ピーク間隔の検出は、すべてのピーク位置を基準にして、実質的に重複する間隔は除外するようにして行うようにしたが、これに限るものではない。例えば、各フレームの任意の1つ以上のピーク位置を基準にしてピーク間隔を検出するようにし、このようにして求めたピーク期間を用いるようにしてもよい。すなわち、すべてのピーク位置を基準位置として用いて、ピーク間隔を検出する必要は必ずしもない。

また、上述した実施の形態においては、1フレームは4秒の期間であって、20Hzのサンプリング周波数のクロック信号を用いるものとして説明したが、これに限るものではない。フレームの時間長、サンプリング周波数は、カーステレオ装置などの機器に搭載されたCPUの性能などに応じて、適宜のものを選択するようにすればよい。

さらに、上述した実施の形態においては、特定したテンポと総音量に応じて、LCDに例えば静止画像を表示するとともに、表示オブジェクトをも表示させて、この表示オブジェクトを移動させるようにしたが、特定したテンポに応じた処理は、これに限るものではない。

例えば、テンポが速い楽曲が再生されている場合には、低域と高域との音域を強調するようにしたり、また、テンポが遅い楽曲が再生されている場合には、サラウンドモードにしたり、リバーブを強めにかけたりするなどの種々の調整を行うようにしてもよい。

つまり、特定したテンポに応じて、イコライザの調整、サラウンドモードの切り換え、音量（ボリューム）の調整等の種々の制御を行うことが可能である。

上述した実施の形態においては、本発明をカーステレオ装置に適用した例を挙げて説明したが、本発明はこれに限るものではない。家庭用ステレオ装置、CD

プレーヤ、MDプレーヤ、DVDプレーヤ、パーソナルコンピュータなどの音声信号を再生して出力するようにする種々のオーディオ装置、オーディオ／ビジュアル装置にこの発明を適用することができる。

本発明を例えば家庭用ステレオ装置に適用した場合には、特定したテンポに応じて、室内の照明の明るさや室温の調整などを行うようにすることも可能である。

また、上述の実施の形態においては、音声信号の帯域分割は、既存の集積回路（IC）を用いて行うものとして説明したが、これに限るものではない。音声信号の帯域分割も例えば、制御部9において実行されるプログラムによって行うようにすることもできる。

本発明は、ソフトウェアによっても十分に実現することができる。これを具体的に示せば、第1番目のプログラムとして、音声信号を処理する装置のコンピュータに、供給される音声信号のレベルが、所定の閾値より大きく、かつ、レベル変化の頂点となっているピーク位置を検出する検出ステップと、所定の単位時間区間において、検出した前記ピーク位置を対象として、少なくとも所定のピーク位置とその他のピーク位置との間の時間間隔を検出する時間間隔検出ステップと、検出した前記時間間隔のうち、発生頻度の多い時間間隔に基づいて、前記音声信号により再生される音声のテンポを特定する特定ステップとを実行するプログラムを作成し、これを有線、無線、あるいは、記録媒体を介して、オーディオ機器やオーディオ／ビジュアル機器に供給し、実行できるようにすることによって本発明に係る装置、方法を実現することもできる。

また、第2番目のプログラムとして、上述の第1番目のプログラムにおいて、特定ステップにおいては、複数の前記単位時間区間において検出されるピーク位置間の時間間隔の発生頻度を累積し、この累積した発生頻度に基づいて、再生される音声のテンポを特定するようにするプログラムを作成することもできる。

また、上述したカーステレオ装置の場合と同様に、第3のプログラムとして、供給される前記音声信号を複数の周波数帯域に分離する帯域分離ステップを設け、検出ステップにおいては、分離された前記複数の周波数帯域の少なくとも1つ以上の帯域毎に前記ピーク位置を検出するようにし、時間間隔検出ステップにおいては、少なくとも1つ以上の帯域毎のピーク位置を対象として、帯域毎に、時間



間隔を検出するようにし、特定ステップにおいては、少なくとも1つ以上の帯域毎に検出される時間間隔のうち、発生頻度の多い時間間隔に基づいて、再生される音声の前記テンポを特定するようにするプログラムを作成することも可能である。

また、第4のプログラムとして、出力しようとする音声信号に基づいて、出力しようとする音声の音量を算出する音量算出ステップと、算出した音量を基準として、ピーク位置を検出する場合に用いる閾値を設定する閾値設定ステップとを設けたプログラムを作成することも可能である。

また、第5のプログラムとして、特定されたテンポに基づいて、メモリに記憶されている画像データの中から画像表示素子に表示する画像の画像データを抽出する画像抽出ステップと、抽出した画像データに応じた画像を画像表示素子に表示する表示ステップとを設けたプログラムを作成することも可能である。

また、第6のプログラムとして、特定された前記テンポに基づいて、画像表示素子に表示する画像の大きさ、移動速度、移動パターンを制御するステップを備えたプログラムを作成することも可能である。

このように、本発明に係るテンポ解析装置及びテンポ解析方法は、プログラムによっても実現可能であり、作成したプログラムは、インターネットや電話網などの種々の電気通信回線やデータ放送によってユーザに提供することが可能であり、また、上述したステップを有するプログラムを記録した記録媒体を配布することによってもユーザに提供することができる。

#### 産業上の利用可能性

上述したように、本発明によれば、自己相関演算などの複雑な演算処理を行うことなく、楽曲などの音声のテンポを簡単かつ正確に検出することができる。また、検出したテンポに応じて情報を提供したり、種々の制御を行ったりするなどのことができるようにされる。ハードウェア割り込みを使用してネットワークの接続されたことを検出し、さらに、リンクを確立させるようにしたので、システムの負荷を最小にすることができるとともに、ネットワークケーブルを接続する

と、ただちにネットワークを使用することができる。

## 請求の範囲

1. 入力音声信号のレベル変化のピークのうち、所定の閾値より大きい複数のピーク的位置を検出するピーク検出手段と、

所定の単位時間区間において、前記ピーク検出手段により検出される前記ピーク位置の間の時間間隔を検出する間隔検出手段と、

前記間隔検出手段により検出される前記時間間隔のうち、発生頻度の多い前記時間間隔に基づいて、前記音声信号により再生される音声のテンポを特定する特定手段と

を備えることを特徴とするテンポ解析装置。

2. 前記特定手段は、複数の前記単位時間区間において検出されるピーク位置間の前記時間間隔の発生頻度を累積し、この累積した発生頻度に基づいて、再生される音声の前記テンポを特定することを特徴とする請求の範囲第1項記載のテンポ解析装置。

3. 前期入力信号を複数の周波数帯域に分離する帯域分離手段を備え、

前記ピーク検出手段は、前記帯域分離手段により分離された複数の帯域のうちの少なくとも1つ以上の帯域毎に前記ピーク位置を検出するものであり、

前記間隔検出手段は、前記ピーク検出手段により検出される少なくとも1つ以上の前記帯域毎の前記ピーク位置の前記時間間隔を検出するものであり、

前記特定手段は、少なくとも1つ以上の前記帯域毎に検出される前記時間間隔のうち、発生頻度の多い時間間隔に基づいて、再生される音声の前記テンポを特定する特徴とする請求の範囲第1項記載のテンポ解析装置。

4. 前記入力音声信号から所定の周波数帯域の音声信号を抽出する帯域抽出手段を備え、

前記ピーク検出手段は、前記帯域抽出手段により抽出された音声信号について前記ピーク位置を検出することを特徴とする請求の範囲第1項記載のテンポ解析装置。

5. 前記入力音声信号の音量を算出する音量算出手段と、

前記音量算出手段により算出された音量を基準として、前記ピーク位置を検出

する場合に用いられる前記閾値を設定する閾値設定手段と

を備えることを特徴とする請求の範囲第1項記載のテンポ解析装置。

6. 前記帯域分離手段により分離された複数の帯域のうちの少なくとも1つ以上の帯域の音声信号の音量を算出する音量算出手段と、

前記音量算出手段により算出された音量を基準として、前記ピーク位置を検出する場合に用いる前記閾値を設定するようにした閾値設定手段と

を備えることを特徴とする請求の範囲第3項記載のテンポ解析装置。

7. 前記帯域抽出手段で抽出された音声信号の音量を算出する音量算出手段と、

前記音量算出手段により算出された音量を基準として、前記ピーク位置を検出する場合に用いる前記閾値を設定する閾値設定手段と

を備えることを特徴とする請求の範囲第4項記載のテンポ解析装置。

8. 画像表示素子と、

前記画像表示素子に表示可能な複数の画像の画像データを記憶する記憶手段と、

前記特定手段により特定される前記テンポに基づいて、前記記憶手段から画像データを選択して読み出し、読み出した前記画像データに応じた画像を前記画像表示素子に表示する表示制御手段と

を備えたことを特徴とする請求の範囲第1項記載のテンポ解析装置。

9. 前記表示手段は、前記記憶手段から読み出す前記画像データに応じた画像を前記画像表示素子に表示する前記画像の大きさ、移動速度、移動パターンの少なくとも1つを制御することを特徴とする請求の範囲第8項記載のテンポ解析装置。

10. 前記表示手段は、前記特定手段により特定される前記テンポと前記音量算出手段により算出された音量とに基づいて、前記記憶手段から画像データを選択して読み出すことを特徴とする請求の範囲第8項記載のテンポ解析装置。

11. 入力音声信号のレベル変化のうち、所定の閾値より大きい複数のピークの位置を検出し、

所定の単位時間区間において、検出した前記ピーク位置の間の時間間隔を検出し、

検出した前記時間間隔のうち、発生頻度の多い時間間隔に基づいて、前記入力音声信号により再生される音声のテンポを特定することを特徴とするテンポ解析

方法。

12. 前記テンポの特定に際しては、複数の前記単位時間区間において検出される前記ピーク位置間の前記時間間隔の発生頻度を累積し、この累積した発生頻度に基づいて、再生される音声の前記テンポを特定することを特徴とする請求の範囲第11項記載のテンポ解析方法。

13. 前記入力音声信号を複数の周波数帯域に分離し、

前記ピーク位置の検出に際しては、分離された前記複数の周波数帯域の少なくとも1つ以上の帯域毎に前記ピーク位置を検出し、

前記時間間隔の検出に際しては、少なくとも1つ以上の前記帯域毎の前記ピーク位置の前記時間間隔を検出し、

前記テンポの特定に際しては、少なくとも1つ以上の前記帯域毎に検出される前記時間間隔のうち、発生頻度の多い時間間隔に基づいて、再生される音声の前記テンポを特定することを特徴とする請求の範囲第11項記載のテンポ解析方法。

14. 前記入力音声信号から所定の周波数帯域の音声信号を抽出し、前記ピーク位置を検出に際しては、前記抽出された音声信号についての前記ピーク位置を検出することを特徴とする請求の範囲第11項記載のテンポ解析方法。

15. 前記入力音声信号の音量を算出し、算出した前記音量を基準として、前記ピーク位置を検出する場合に用いる前記閾値を設定することを特徴とする請求の範囲第11項記載のテンポ解析方法。

16. 前記分離された複数の帯域のうちの少なくとも1つ以上の帯域の音声信号の音量を算出し、算出した前記音量を基準として、前記ピーク位置を検出する場合に用いる前記閾値を設定することを特徴とする請求の範囲第13項記載のテンポ解析方法。

17. 前記算出された前記音量の音量を算出し、算出した前記音量を基準として、前記ピーク位置を検出する場合に用いる前記閾値を設定することを特徴とする請求の範囲第14項記載のテンポ解析方法。

18. 特定された前記テンポに基づいて、記憶手段に記憶されている複数の画像データの中から画像データを選択して読み出し、読み出した前記画像データに応じた画像を前記画像表示素子に表示することを特徴とする請求の範囲第11項記

載のテンポ解析方法。

19. 特定された前記テンポに基づいて、画像表示素子に表示する画像の大きさ、移動速度、移動パターンを制御することを特徴とする請求の範囲第18項記載のテンポ解析方法。

20. 特定された前記テンポと算出された前記音量とに基づいて、前記記憶手段に記憶されている複数の画像データを選択して読み出すことを特徴とする請求の範囲第18項記載のテンポ解析方法。

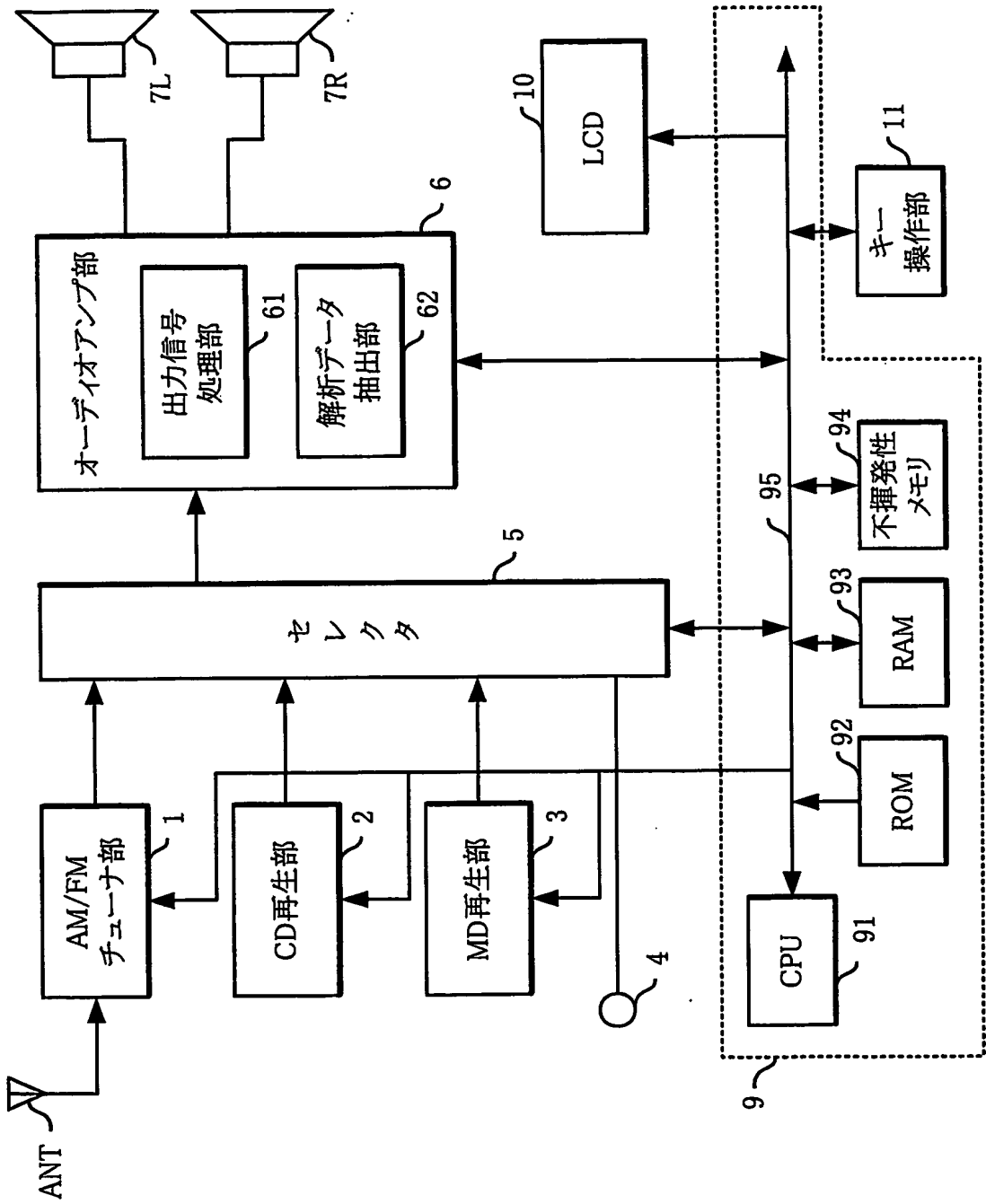


FIG.1

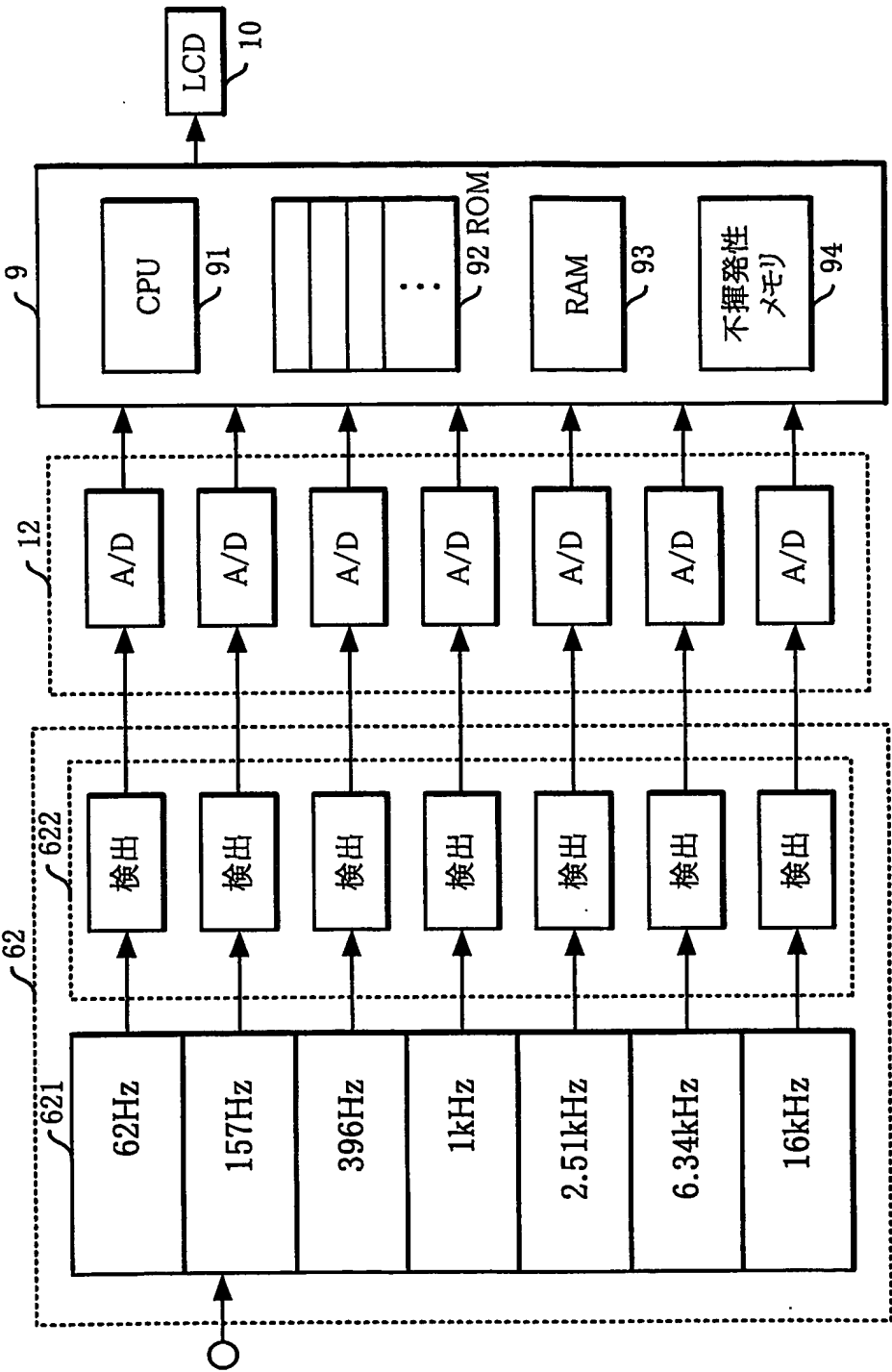


FIG.2



3/14

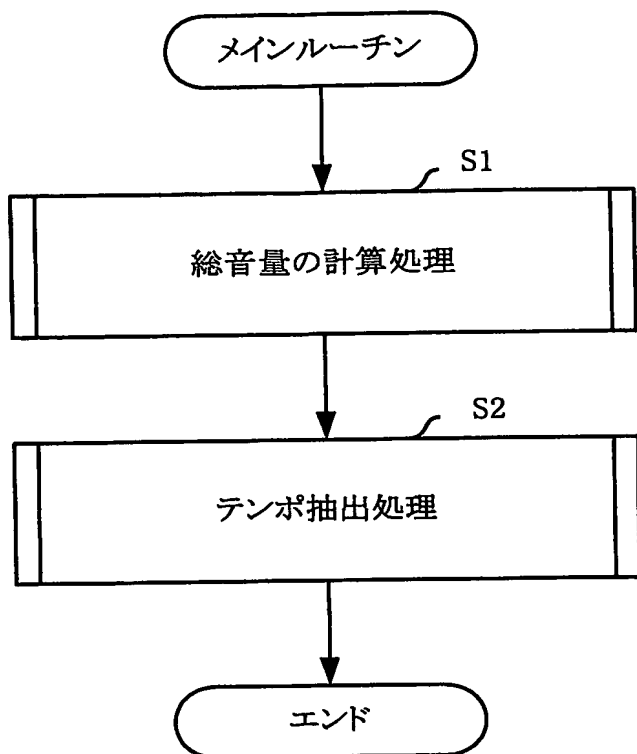


FIG.3

4/14

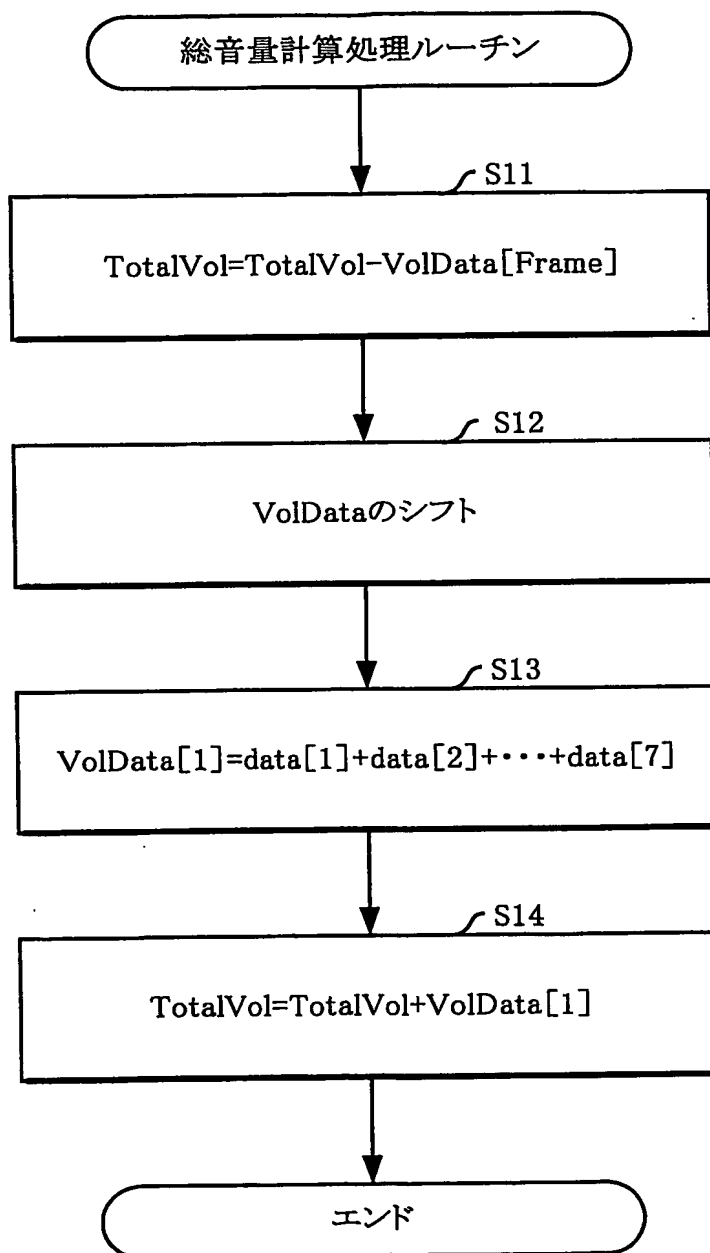


FIG.4

5/14

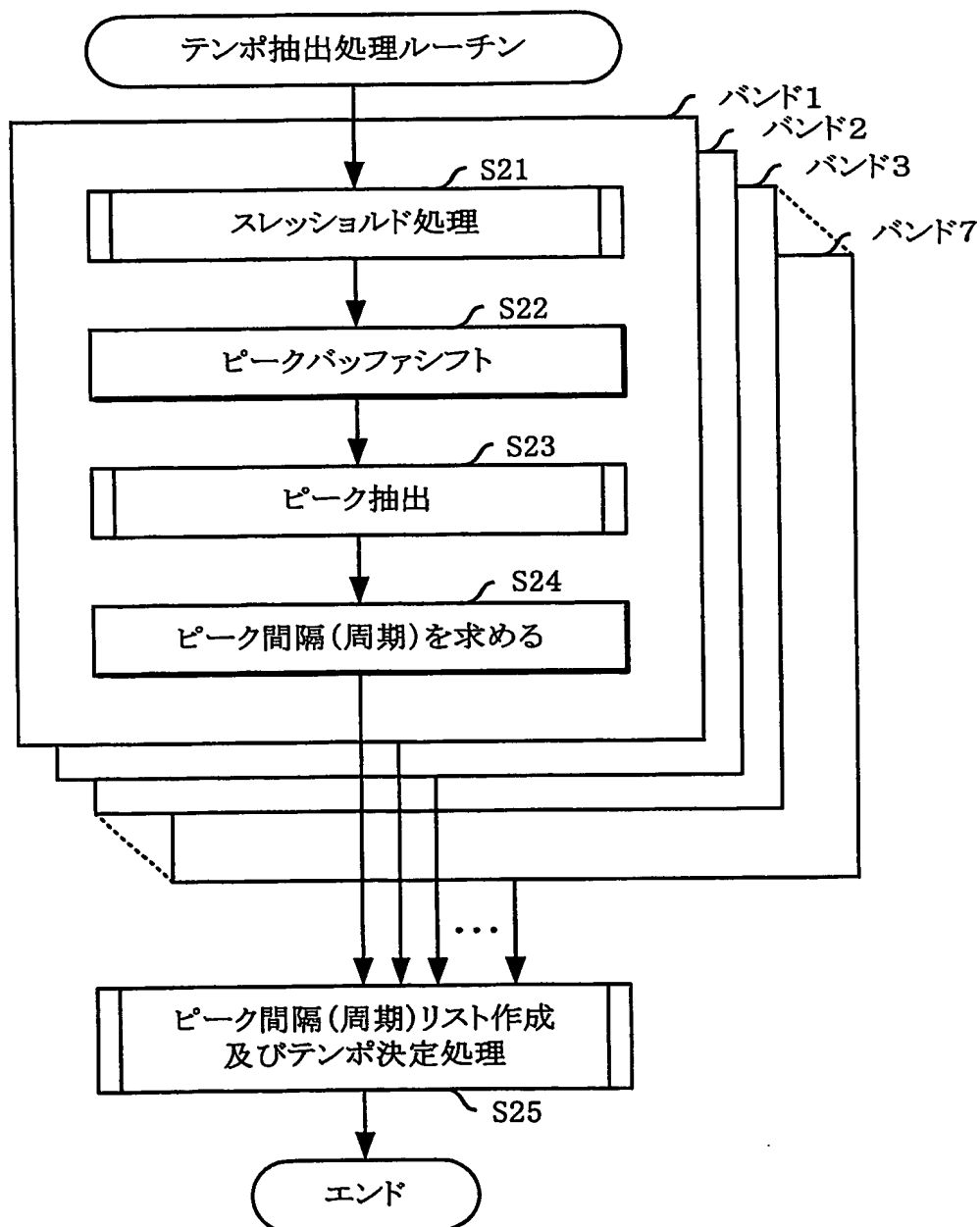


FIG.5

6/14

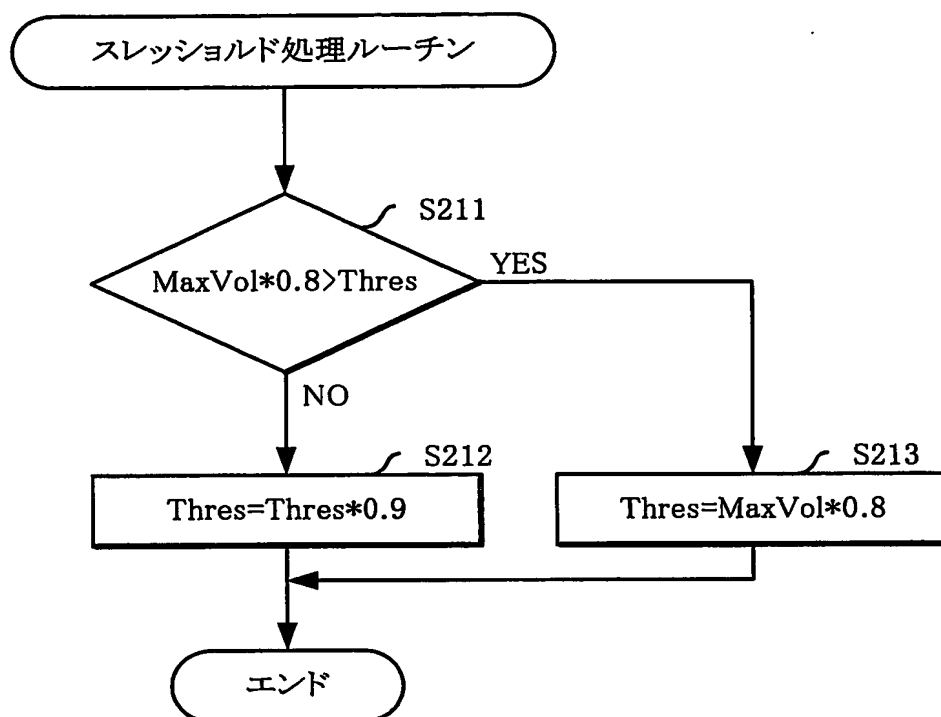


FIG.6

7/14

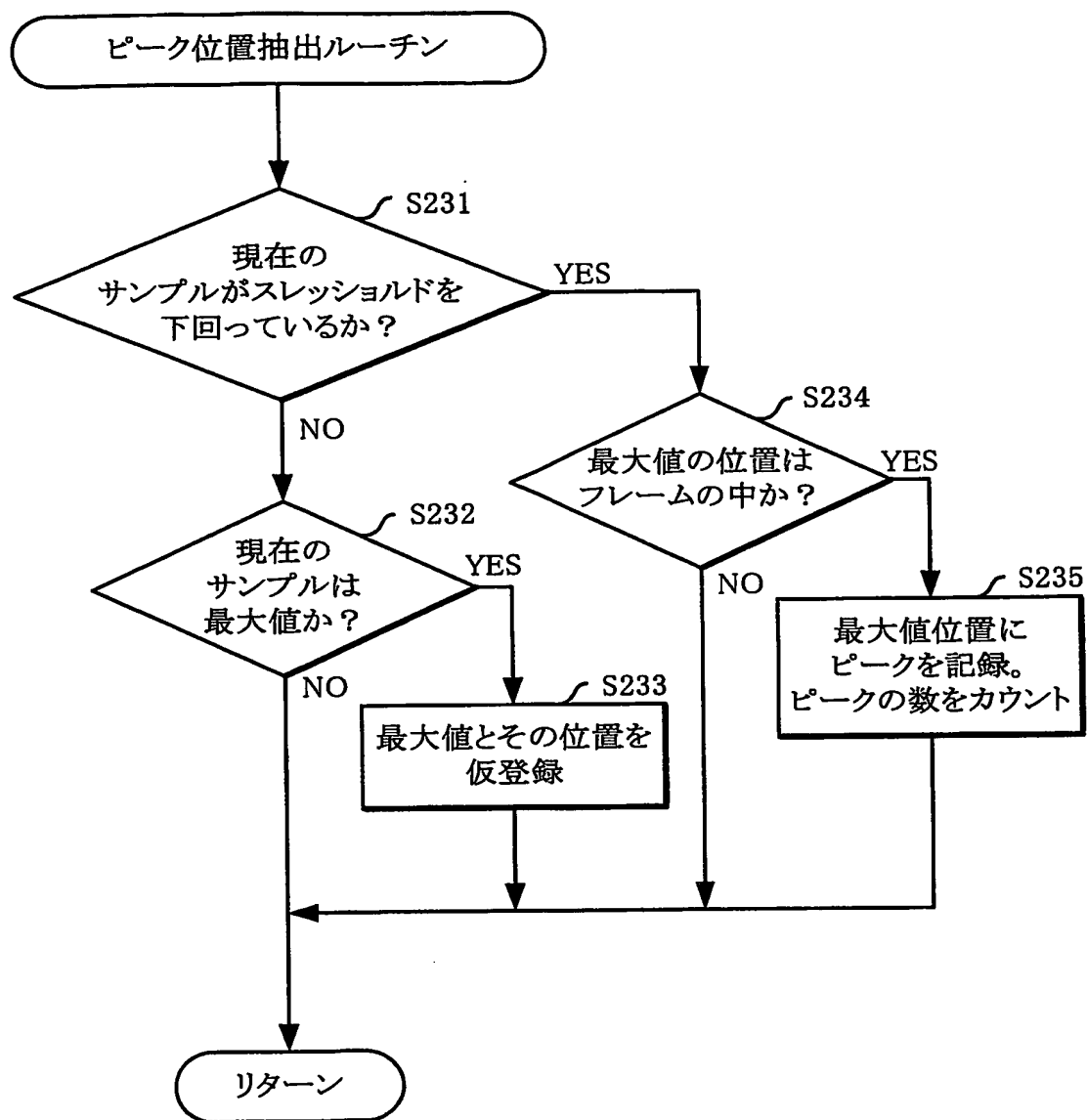


FIG.7

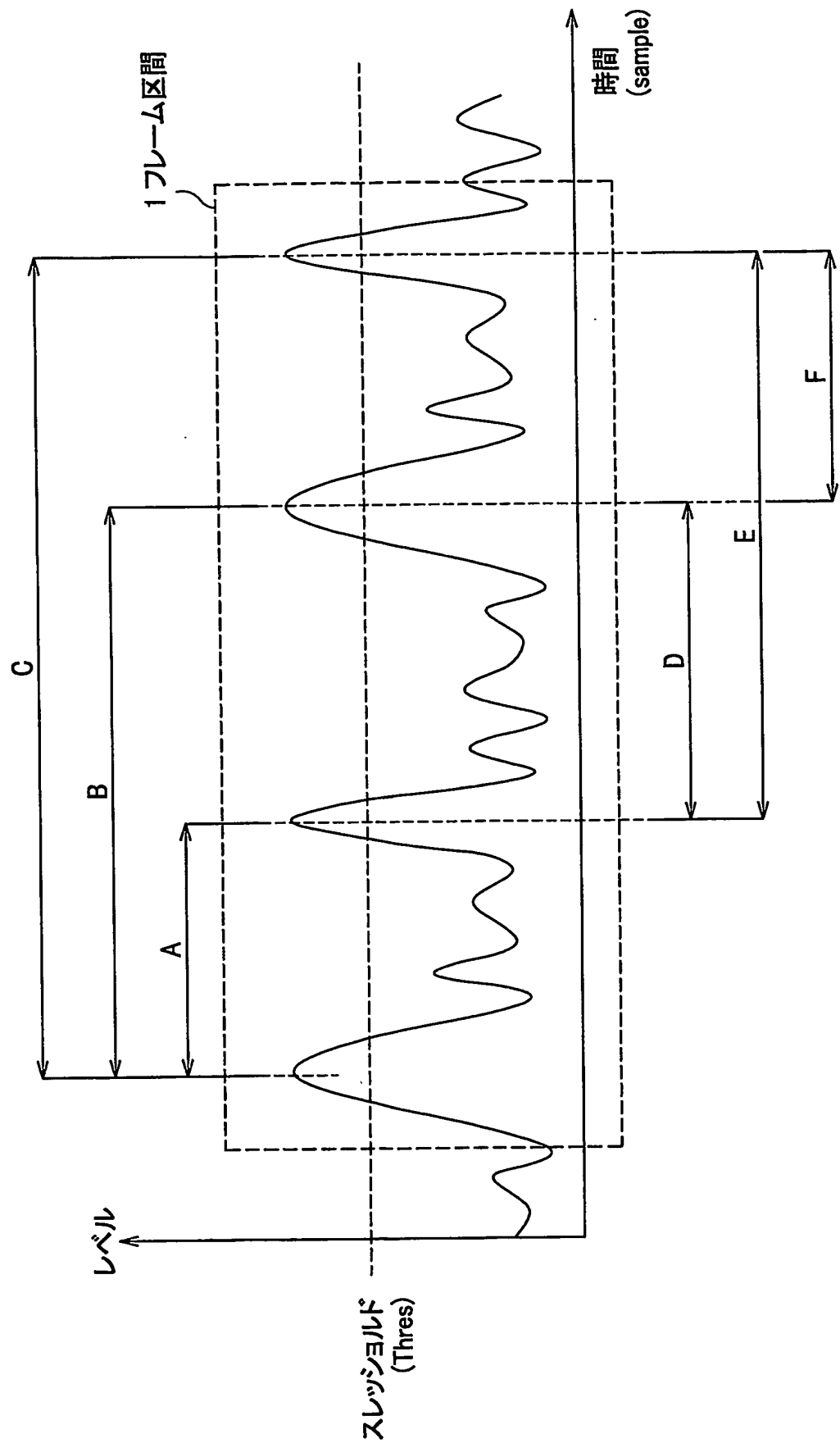


FIG.8

9/14

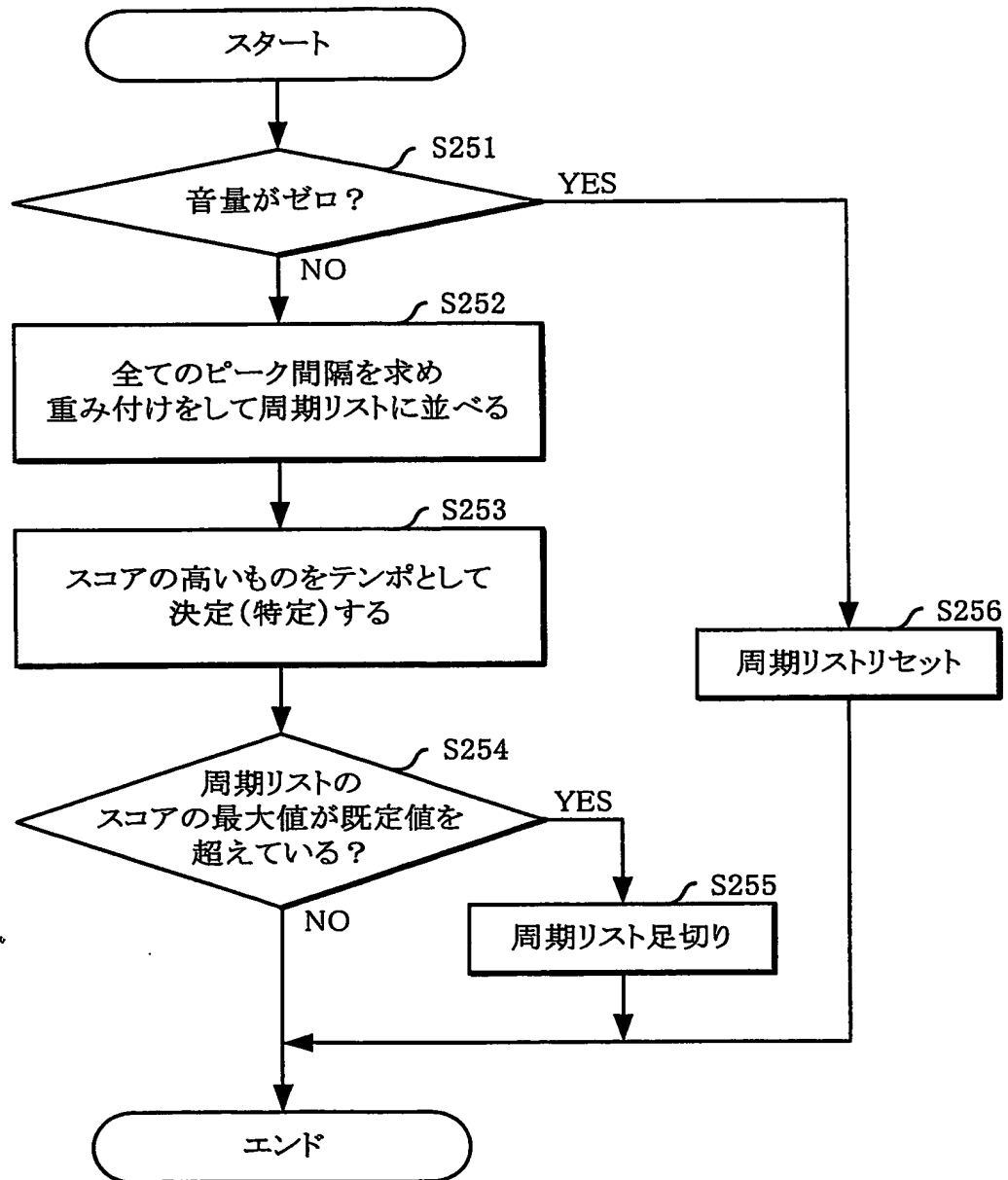


FIG.9

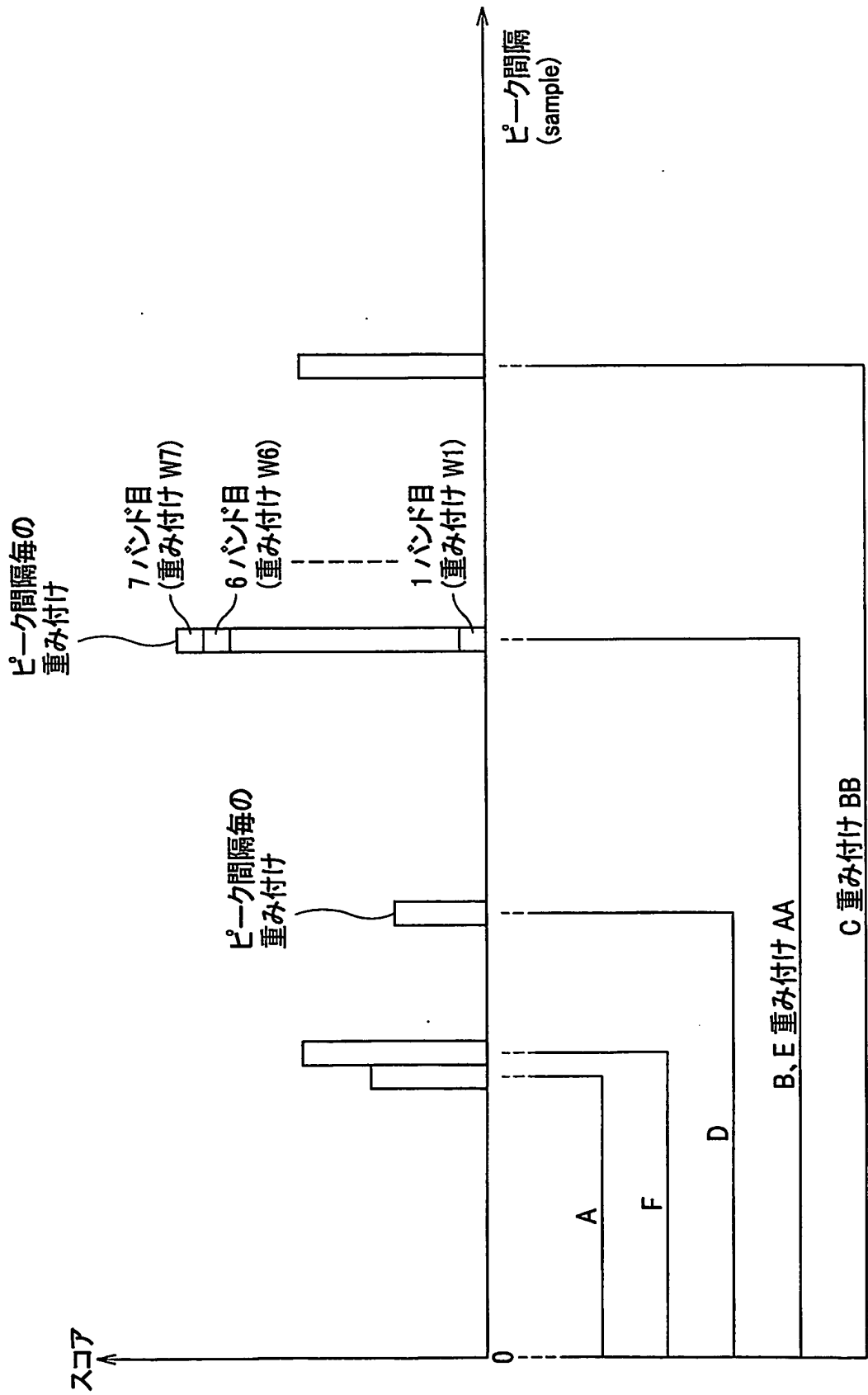


FIG.10



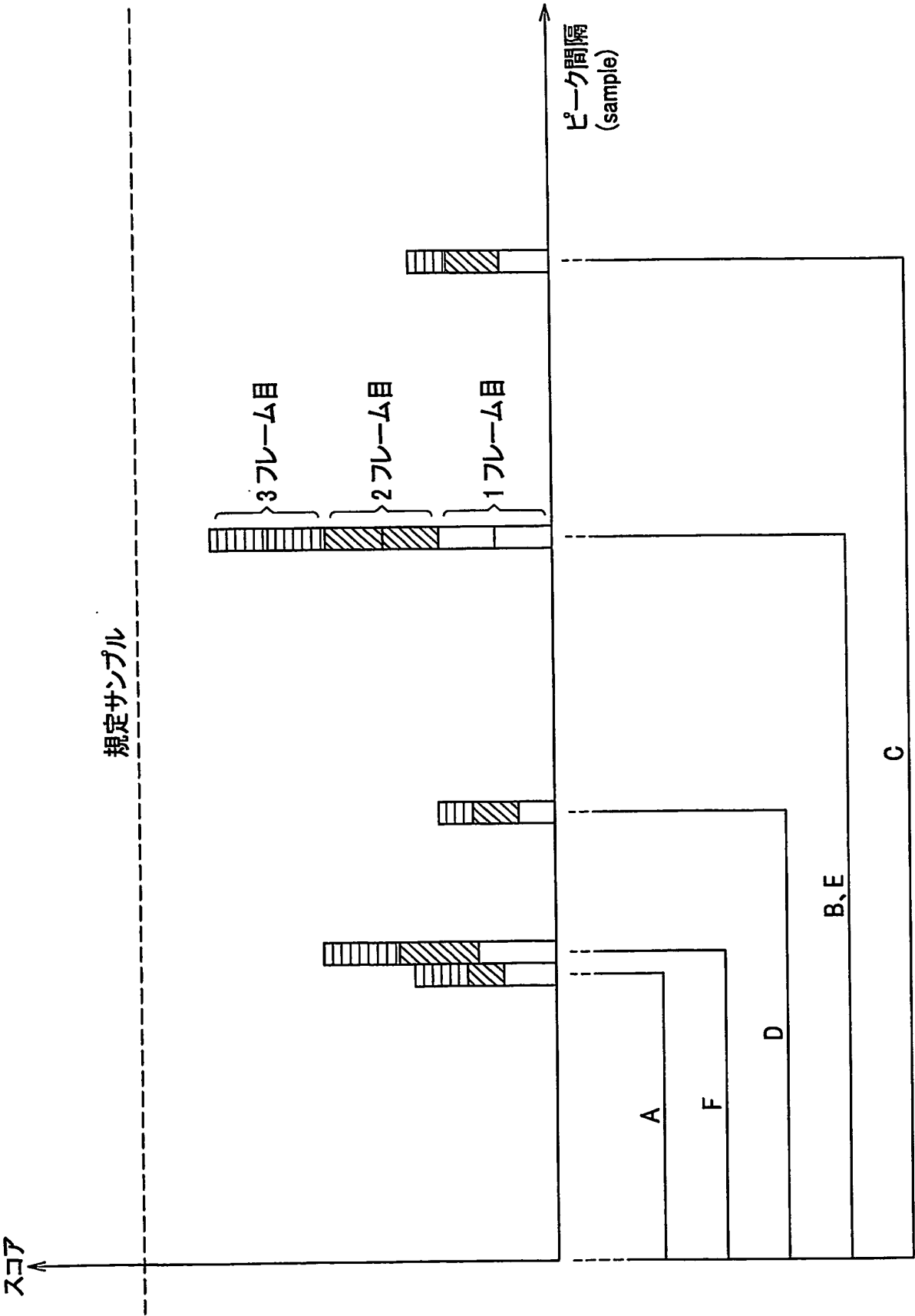


FIG.11

12/14

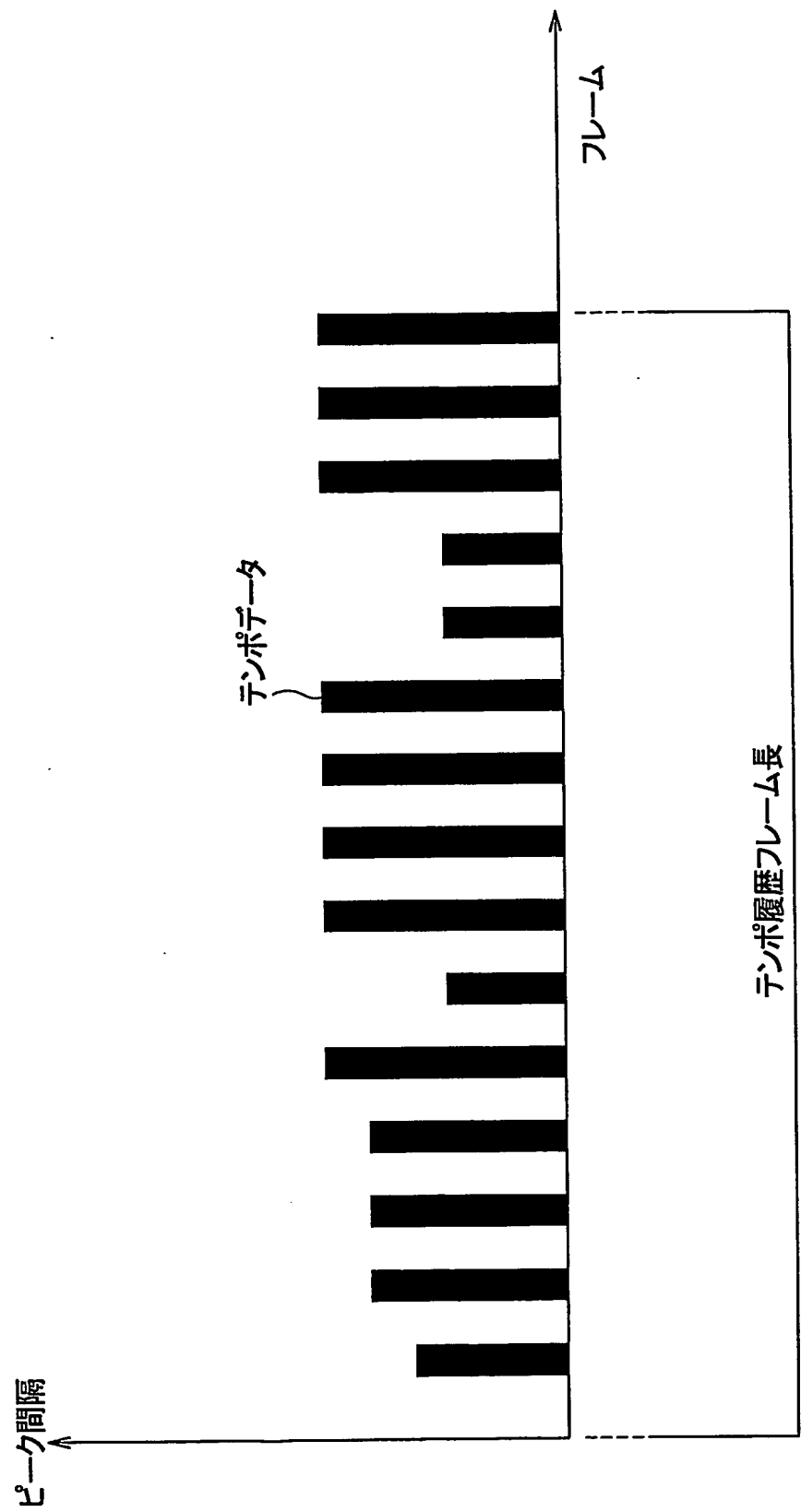


FIG.12

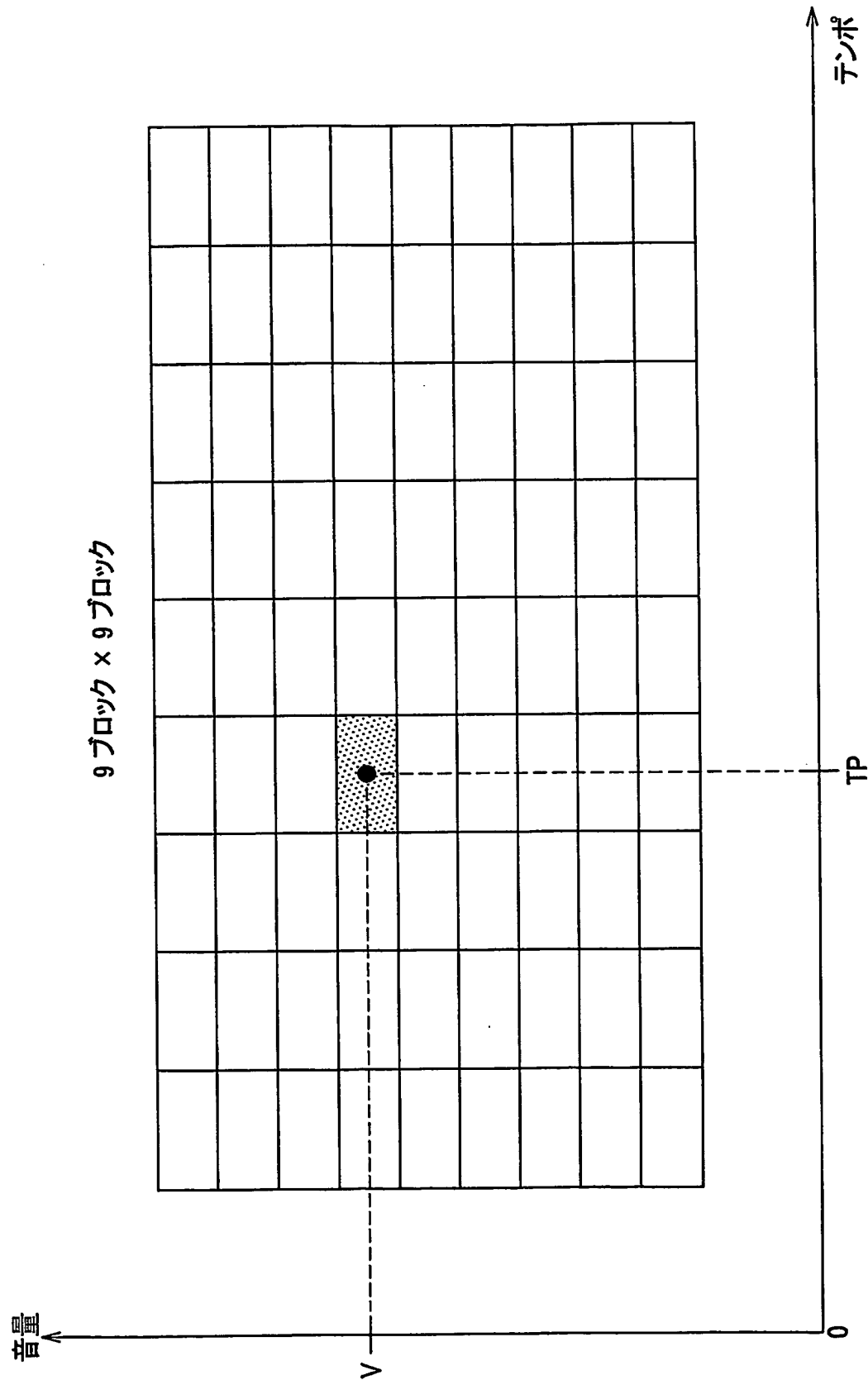


FIG.13

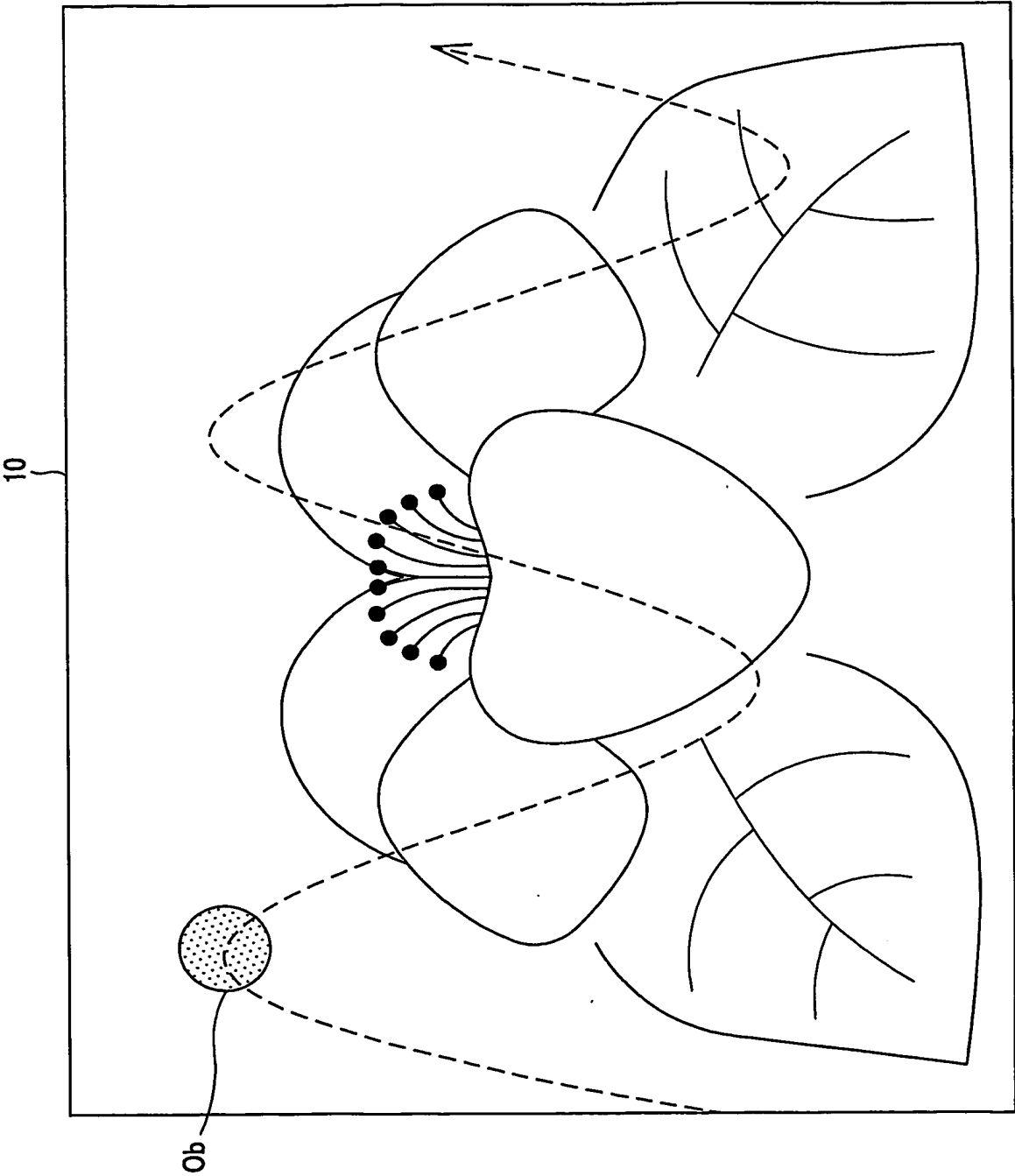


FIG.14

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/003010

**A. CLASSIFICATION OF SUBJECT MATTER**  
Int.Cl<sup>7</sup> G10H1/40, G10G3/04

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
Int.Cl<sup>7</sup> G10H1/40, G10G3/04

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  
Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2004  
Kokai Jitsuyo Shinan Koho 1971-2004 Toroku Jitsuyo Shinan Koho 1994-2004

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2002-207482 A (Matsushita Electric Industrial Co., Ltd.), 26 July, 2002 (26.07.02), Par. Nos. [0030] to [0083]; Figs. 6, 9 (Family: none)	1-20
Y	JP 2000-267654 A (Aiwa Co., Ltd.), 29 September, 2000 (29.09.00), Full text; all drawings (Family: none)	1-20
Y	JP 8-76777 A (Yamaha Corp.), 22 March, 1996 (22.03.96), Column 16; lines 1 to 4 & US 5863206 A Column 10, lines 55 to 58	8-10, 18-20

☒ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search  
24 May, 2004 (24.05.04)

Date of mailing of the international search report  
08 June, 2004 (08.06.04)

Name and mailing address of the ISA/  
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/003010

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2000-311251 A (Toshiba Corp.), 07 November, 2000 (07.11.00), Par. Nos. [0084] to [0086] & EP 1031945 A2 Column 11, lines 15 to 32	8-10, 18-20
Y	JP 10-319957 A (Enix Corp.), 04 December, 1998 (04.12.98), Par. Nos. [0008] to [0009], [0029] to [0030] (Family: none)	8-10, 18-20
Y	JP 2000-276137 A (Yamaha Corp.), 06 October, 2000 (06.10.00), Par. Nos. [0007] to [0018] & US 6225545 B1 Column 3, line 30 to column 6, line 42	9, 10, 19, 20

<b>A. 発明の属する分野の分類 (国際特許分類 (IPC))</b> Int. Cl <sup>7</sup> G10H1/40, G10G3/04			
<b>B. 調査を行った分野</b> 調査を行った最小限資料 (国際特許分類 (IPC)) Int. Cl <sup>7</sup> G10H1/40, G10G3/04			
最小限資料以外の資料で調査を行った分野に含まれるもの 日本国実用新案公報 1922-1996年 日本国公開実用新案公報 1971-2004年 日本国実用新案登録公報 1996-2004年 日本国登録実用新案公報 1994-2004年			
国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)			
<b>C. 関連すると認められる文献</b>			
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号	
Y	JP 2002-207482 A (松下電器産業株式会社) 2002. 07. 26, 段落【0030】-【0083】, 第6図, 第9図 (ファミリーなし)	1-20	
Y	JP 2000-267654 A (アイワ株式会社) 2000. 09. 29, 全文, 全図 (ファミリーなし)	1-20	
Y	JP 8-76777 A (ヤマハ株式会社) 1996. 03. 22, 第16欄第1行-第4行 & US 5863206 A, 第10欄第 55行-第58行	8-10, 18-20	
<input checked="" type="checkbox"/> C欄の続きにも文献が列挙されている。 <input type="checkbox"/> パテントファミリーに関する別紙を参照。			
* 引用文献のカテゴリー 「A」 特に関連のある文献ではなく、一般的技術水準を示すもの 「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの 「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す) 「O」 口頭による開示、使用、展示等に言及する文献 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願			
の日の後に公表された文献 「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの 「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの 「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの 「&」 同一パテントファミリー文献			
国際調査を完了した日 24. 05. 2004		国際調査報告の発送日 08. 6. 2004	
国際調査機関の名称及びあて先 日本国特許庁 (ISA/JP) 郵便番号100-8915 東京都千代田区霞が関三丁目4番3号		特許庁審査官 (権限のある職員) 小宮 慎司 5 C 9567 電話番号 03-3581-1101 内線 3539	

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	JP 2000-311251 A (株式会社東芝) 2000. 11. 07, 段落【0084】－【0086】 & EP 1031945 A2, 第11欄第15行－第32行	8-10, 18-20
Y	JP 10-319957 A (株式会社エニックス) 1998. 12. 04, 段落【0008】－【0009】, 【0029】－【0030】 (ファミリーなし)	8-10, 18-20
Y	JP 2000-276137 A (ヤマハ株式会社) 2000. 10. 06, 段落【0007】－【0018】 & US 6225545 B1, 第3欄第30行－第6欄第42行	9, 10, 19, 20